INVESTIGACION y CIENCIA

LA OLA DE LOS FACTORES DE TRANSCRIPCION

LASERES EN CIRUGIA

POR QUE UN CUASAR ECLIPSA MIL GALAXIAS

Edición española de
SCIENTIFIC AMERICAN



EL ARCO PRIMITIVO

AGOSTO 1991
600 PTAS.



Agosto de 1991 Número 179

Edición española de SCIENTIFIC AMERICAN

SECCIONES

COLABORADORES DE ESTE NUMERO

Asesoramiento y traducción:

Fernando Pérez Aragüés: *Arthur Stanley Eddington*; Ramón Pascual: *El cuásar 3C 273*; Santiago Torres: *La proteína M de los estreptococos*; Manuel Puigcerver: *Nubes estratosféricas polares y empobrecimiento en ozono*; J. Vilardell: *Diseño y construcción de los arcos primitivos* y *Hace...*; Ana Mª García: *Cirugía con láser*; Consuelo García Majonero: *Mirando hacia dentro*; Luis Bou: *Juegos matemáticos*.

Ciencia y sociedad:

Joandomènec Ros y J. Enric Llebot

Ciencia y empresa:

Manuel Puigcerver

Libros:

Luis Mas, Marià Alemany, Antonio F. Costa, Luis Alonso y Domingo Fletcher

PROCEDENCIA DE LAS ILUSTRACIONES

Página	Fuente
5	Biblioteca de la Real Sociedad Astronómica
6-7	Johnny Johnson
8	Biblioteca de la Real Sociedad Astronómica
9	E. Eddington
11	Observatorio J.-L. Nieto de Toulouse
12	George Retseck
13-14	Edward Bell
15	George Retseck
16	Laurie Grace
17	Edward Bell
19	Vincent A. Fischetti
20	Tomo Narashima
21	Vincent A. Fischetti
22	Johnny Johnson
23	Tomo Narashima
24	Edward Bell
25	Tomo Narashima
34-35	NASA
36-40	Ian Worpole
43-51	Antonio Celada y Silvia Nuere
52	Charles E. Grayson
54-57	Hank Iken
58	Charles E. Grayson
61	Peter Hering
62	Andrew Christie; Michael W. Berns (*fotografía*)
63	J. Stuart Nelson, Instituto Beckman del Láser, Universidad de California, Irvine (*arriba*), Marjorie Mosier, Instituto Beckman del Láser (*abajo*)
64	Richard Straight
65	Andrew Christie
66	Michael W. Berns
67	Aaron Lewis y Daniel Palanker
68-69	George V. Kelvin
70-71	Gary A. Glatzmaier, Laboratorio Nacional de los Alamos (*izquierda*), Adam M. Dziewonski, Universidad de Harvard y John H. Woodhouse, Universidad de Oxford (*derecha*)
72	Jason Küffer
74	Johnny Johnson (*arriba*), Adam M. Dziewonski (*abajo*)
75	Jeremy Bloxham, Universidad de Harvard y David Gubbins, Universidad de Leeds
76	Jason Küffer
77	Philippe Machetel y Patrice Weber, Grupo de Investigación de Geodesia Espacial (*izquierda*), Johnny Johnson (*derecha*)
78-79	Andrew Christie
80	Johnny Johnson



EN LA ILUSTRACION de la portada se muestra un arco compuesto del siglo XI para caza y tiro al vuelo construido en la India. Revestido de corteza de árbol primorosamente decorada, este arco está compuesto de madera, nervios animales y cuerno, combinación de materiales que produce un arma aún más poderosa que el arco medieval inglés. (Véase "Diseño y construcción de los arcos primitivos", por Edward McEwen, Robert L. Miller y Christopher A. Bergman, en este mismo número.)

INVESTIGACION Y CIENCIA

DIRECTOR GENERAL Francisco Gracia Guillén
DIRECTOR EDITORIAL José María Valderas Gallardo
DIRECTORA DE ADMINISTRACIÓN Pilar Bronchal Garfella
PRODUCCIÓN César Redondo Zayas
SECRETARÍA Purificación Mayoral Martínez
EDITA Prensa Científica, S. A. Viladomat, 291 6º 1ª - 08029 Barcelona (ESPAÑA)
Teléfonos 321 81 91 - 321 83 48 Telefax 419 47 82

SCIENTIFIC AMERICAN

EDITOR Jonathan Piel
BOARD OF EDITORS Alan Hall, *Executive Editor*; Michelle Press, *Managing Editor*; Timothy M. Beardsley; Elizabeth Corcoran; Deborah Erickson; Marguerite Holloway; John Horgan; Philip Morrison, *Book Editor*; Corey S. Powell; John Rennie; Philip E. Ross; Ricki L. Rusting; Russell Ruthen, Gary Stix; Paul Wallich; Philip M. Yam.
PUBLISHER John J. Moeling, Jr.
ADVERTISING DIRECTOR Robert F. Gregory
PRESIDENT AND CHIEF EXECUTIVE OFFICER Claus-Gerhard Firchow
CHAIRMAN OF THE BOARD Dr. Pierre Gerckens
CHAIRMAN EMERITUS Gerard Piel

SUSCRIPCIONES

Prensa Científica S. A.
Viladomat, 291 6º 1ª
08029 Barcelona (España)
Teléfonos 321 81 91 - 321 83 48

Precios de suscripción, en pesetas:

	Un año	Dos años
España	6600	12.000
Extranjero	7300	13.400

Ejemplares sueltos:

Ordinario: 600 pesetas
Extraordinario: 775 pesetas

— Todos los precios indicados incluyen el IVA, cuando es aplicable.

— En Canarias, Ceuta y Melilla los precios incluyen el transporte aéreo.

— El precio de los ejemplares atrasados es el mismo que el de los actuales.

DISTRIBUCION

para España:

MIDESA
Carretera de Irún, km. 13,350
(Variante de Fuencarral)
28049 Madrid Tel. 652 42 00

para los restantes países:

Prensa Científica, S. A.
Viladomat, 291 6º 1ª - 08029 Barcelona
Teléfonos 321 81 91 - 321 83 48

PUBLICIDAD

Madrid: Gustavo Martínez Ovín
Menorca, 8, bajo, centro, izquierda.
28009 Madrid
Tel. 409 70 45 - Fax 409 70 46

Cataluña: Marcel Klein
M. K. Publicidad
Ortigosa, 14-16, 3º, D. 20
08003 Barcelona
Tel. 268 45 05 - Fax 268 16 07

ISSN 0210-136X Dep. legal: B. 38.999-76

Fotocomposición: Tecfa. Línea Fotocomposición, S.A. Almogàvers, 189 - 08018 Barcelona
Fotocromos reproducidos por Tecfa. Línea Fotomecánica, Almogàvers, 189 - 08018 Barcelona
Imprime Rotographik, S.A. Ctra. de Caldes, km 3,7 - Santa Perpètua de la Mogoda (Barcelona)

Printed in Spain - Impreso en España

Hace...

...cincuenta años

SCIENTIFIC AMERICAN: "Si se descubre un sistema para diezmar las filas enemigas mediante enfermedades contagiosas o venenos mortales, transportados hasta territorio enemigo, este 'cruel' e 'inhumano' método de lucha será aceptado por todos, pese a todos los reparos emocionales. Por suerte para la raza humana, carecen de fundamento los miedos exagerados hacia los efectos devastadores de la guerra bacteriológica. Una cucharada de la toxina de *Bacillus botulinus* bastaría para envenenar a la población de una ciudad grande. Un avión podría transportar suficiente toxina del botulismo para destruir la población mundial. Pero ello no es sencillo en la práctica. Si bien no sería difícil fabricar la cantidad necesaria de botulina y transportarla, el problema real es cómo administrarla. Durante la primera guerra mundial se fabricaron más de 100.000 millones de balas, suficientes para matar 50 veces a todos los habitantes del planeta; pero aún quedamos algunos con vida."

"El Doctor William Bromme, de Detroit, ha declarado que un nuevo tratamiento, eficaz en el 100 por 100 de los casos, promete el control total de la gonorrea. Según el galeno, en tres días de tratamiento con sulfatiazol, sanaron cien enfermos sobre cien. Todos los métodos anteriores fracasaban porque los fármacos no podían llegar a los gérmenes, que rápidamente se hundían bajo la superficie de la piel hasta los tejidos más profundos, donde habitan los gonococos. Predice el Doctor Bromme que, dentro de 20 años, no habrá casos de secuelas gonorreicas u otras complicaciones para mostrar a los estudiantes de medicina."

"El paso siguiente del experimento fue averiguar que ocurriría si se dejaran mudos a los murciélagos. Los biólogos de Harvard Galambos y Griffin pudieron amordazar temporalmente a uno de esos animales, atándole primero el morro con hilo de lencería y sellándole luego los labios con colodión. El resultado fue que el porcentaje de choques subió a 65%. El emisor ultrasónico, o sea, el murciélago, ya no podía lanzar descargas sónicas en vuelo. Para asegurarse doblemente de que los ultrasonidos eran realmente gritos que nacían de su laringe los investigadores hundían en agua la cabeza del animal a lo largo de intervalos cortos. Durante éstos, de la grabadora no salían sonidos de alta frecuencia."

"Que las modernas carrocerías de acero de los automóviles son unas pantallas muy eficaces contra los rayos fue demostrado hace poco en los laboratorios de la Westinghouse Electric and Manufacturing Company, donde se lanzó una descarga artificial de 3.000.000 V contra el techo de un vehículo. Aunque el rayo incidió sobre el techo del automóvil a pocos centímetros del ingeniero sentado en su interior, el único rastro de daños fueron unas pequeñas quemaduras en el metal."

...cien años

SCIENTIFIC AMERICAN: "La naturaleza tropical se presta bien a las artes del envenenador. Por doquier abundan los productos vegetales de los que pueden extraerse los venenos más mortíferos por procedimientos sencillos. Sustancias como la lobelia, la nuez vómica, la belladona, los ácidos prúsico y oxálico, la uriquitina y el manzanillo, amén de otras muchas sustancias menos conocidas abundan por doquier. En el reino animal se conoce perfectamente la potencia de la sangre pútrida, el veneno de las tarántulas, los escorpiones, los ciempiés y, menos, de las serpientes."



Un modo seguro de manejar una boa constrictor.

"Chicago es una ciudad emprendedora hasta lo pasmoso y, en más de un aspecto, sin par. Ganó la sede de la Feria Mundial frente a todas sus competidoras y ahora cuenta con una mujer ingeniero. Un compañero suyo afirma que no le perdonaron fácilmente que fuera mujer; su examen, nos cuenta el articulista, fue algo más riguroso que lo usual. Entre otras cosas, le preguntaron cuál es el tamaño de la válvula de descarga de un motor de siete caballos de potencia y qué haría si la válvula se pegase. Al final de la prueba, los examinadores escribieron "aprobado" en su ejercicio y la señorita De Barr es ya un ingeniero diplomado en máquinas de vapor con todas las de la ley."

"Los obreros del siglo XIX trabajaban por encima de sus fuerzas y con el fin de mantenerlas recurren a estimulantes: café, té, especias, alcohol o tabaco. Estos sobreexcitan los nervios, con la secuela del insomnio y, para vencer a éste, recurren a los narcóticos, a cuya cabeza figura la morfina. Lo que origina el peligro de caer en el hábito de estas drogas es la cobardía y degeneracion de nuestros tiempos. Nadie ha de sufrir dolor, no importa su brevedad. No hay que extraer ni una muela, ningún niño puede venir al mundo sin que se emplee un anodino y, cuando llega la muerte, hay que provocar la eutanasia."

"El plátano pertenece a la familia de las liliáceas y es una azucena tropical desarrollada de la que, tras siglos de cultivo, se han eliminado las semillas y el fruto. Acerca de las cualidades del cultivo de esta fruta, Humboldt dijo que, en la tierra donde crecieran 100 kilos de patata, también crecería 36 kilos de trigo, pero también 4500 kilos de plátanos."

"Mr. G. R. O'Reilly ha viajado mucho en busca de reptiles y serpientes, y tiene gran práctica en capturar vivas las más temidas. En la ilustración se muestra de qué modo maneja una boa constrictor en el Museo del Central Park."

Arthur Stanley Eddington

Dirigió una expedición astronómica para comprobar la teoría de la relatividad general de Einstein, defendió la idea de un universo en expansión y describió la estructura interna de las estrellas

Sir William McCrea

Es difícil explicar al lector de hoy el aura de respeto que rodeaba a Sir Arthur Stanley Eddington en el período entreguerras. Ejerció una gran influencia en el desarrollo del pensamiento científico. En primer lugar, por sus propias contribuciones a la astronomía y la astrofísica y, así mismo, por su aportación al trabajo de otros científicos de su época. En efecto, a menudo parecía capaz de captar las implicaciones de un descubrimiento con más profundidad que su autor y, además, sabía exponerlo con mayor facilidad. Por último, sus intentos de explicar los fundamentos de la física han influido en el pensamiento de los científicos de las generaciones siguientes, aunque muchos de ellos nunca supieran que se habían embarcado en una búsqueda que Eddington comenzó.

En el segundo y tercer decenio de nuestro siglo destacaron personalidades singulares en distintos dominios: Churchill y Roosevelt en política, Shaw y Eliot en literatura y Bohr, Einstein, Rutherford y Eddington en física, por citar unos cuantos. Hoy no se repite ese fenómeno. No queremos afirmar que no existan talentos semejantes, sino que, por razones obvias, es difícil que éstos sean apreciados de la misma forma. En el ámbito científico, la búsqueda del conocimiento se ha transferido del individuo al equipo y, en cierta manera, de los creadores a los divulgadores de ideas. En este sentido Eddington fue, sin duda, el pionero de los divulgadores, pero, por encima de todo, fue creador y desde esta perspectiva hemos redactado nuestro artículo.

Bohr, Einstein y Rutherford hicieron grandes descubrimientos, pero desempeñaron un papel secundario en la divulgación de su trabajo más allá de los círculos profesionales. Eddington tenía otro talante: algunos científicos asistieron a sus clases magistrales, muchos más leyeron algún trabajo suyo y casi todos tuvieron la convicción de haberle comprendido. Parecía disfrutar de una especial capacidad para penetrar en los secretos del universo y, para algunos, su obra escrita raya en la perfección.

Por supuesto, el valor último de su trabajo no descansa en el porte casi místico que poseía en el momento cumbre de su carrera, pero su carismático aspecto reclamaba la atención hacia sus ideas. Este atractivo personal provenía principalmente de la gran fe que depositaba en su propia intuición, que, por otra parte, tantas veces había resultado justificada.

Eddington nació en 1882, de padres cuáqueros. Perdió a su padre dos años más tarde. Estudió matemáticas y física en el Owens College (hoy Universidad de Manchester) y en la Universidad de Cambridge, donde llegó a ser "senior wrangler", la calificación más alta que se podía allí recibir en matemáticas. Trabajó en el Real Observatorio de Greenwich desde 1906 hasta 1913, año en que se hizo cargo de la cátedra Plumian de astronomía y filosofía experimental de Cambridge. Seis años después encabezó una famosa expedición a una isla próxima a Africa ecuatorial para observar un eclipse total que determinaría si la teoría general de la relatividad de Einstein era o no correcta. Presidió la Real Sociedad Astronómica, se le concedió el título de Sir en 1930 y murió en 1944.

Hasta caer víctima de la que sería su última enfermedad, Eddington fue un hombre vigoroso. Le gustaba nadar y jugar al golf (no demasiado bien, por lo que se sabe). Ciclista destacado, llevaba un cuidadoso registro de sus salidas más largas, a partir de las cuales evaluaba su parámetro n, es decir, el número de días en los que había recorrido n o más kilómetros en su bicicleta. El último valor alcanzado fue de 120.

Terriblemente tímido en encuentros sociales intrascendentes y poco hábil al dar una clase o en una conversación improvisada, poseía una facilidad especial para dictar conferencias y era un maestro presentando un conciso resumen, sobre un artículo de investigación, en una reunión científica. Quienes le conocían bien veían en él un compañero atento y sociable. Baste un ejemplo: cuando su colega y amigo, el astrónomo F.J.M. Stratton, cayó, ya en su madurez, víctima de una grave enfermedad que amenazaba con dejarlo sumido en una profunda depresión, Eddington inventaba, casi a diario, un nuevo acertijo intelectual para mantener a Stratton alerta y acelerar su recuperación.

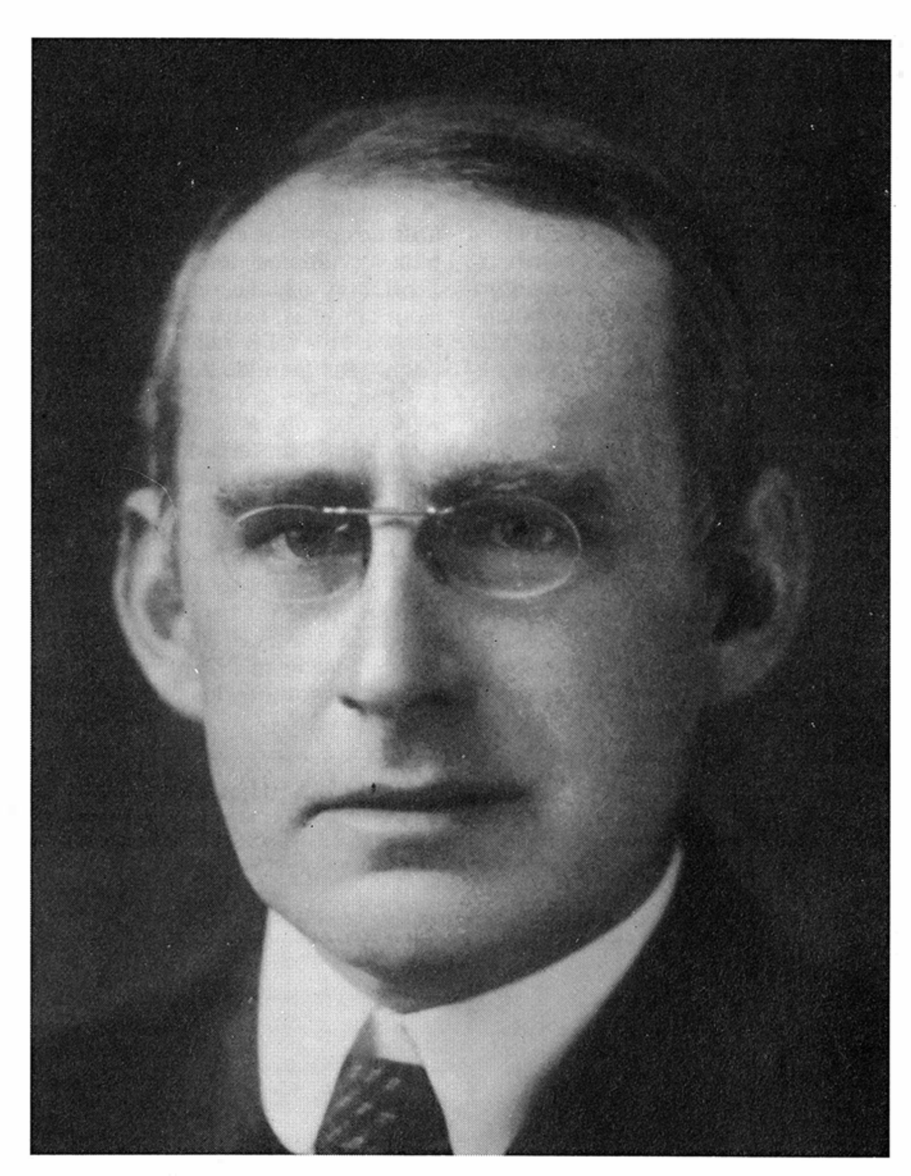
Su educación religiosa, cuáquera, conformó la actitud que mantuvo ante la vida y, por tanto, podría parecer inevitable que él continuara fiel a esta idea. Sucede a menudo, sin embargo, todo lo contrario: muchas personas de mentalidad independiente reaccionan en contra de la educación recibida. Eddington siguió, pues, en su creencia de manera consciente.

SIR WILLIAM McCREA es profesor emérito de astronomía teórica en la Universidad de Sussex, antiguo presidente de la Real Sociedad Astronómica (cargo que Eddington también desempeñó) y miembro de la Regia Sociedad Londinense. Se le concedió el título de Sir por sus aportaciones a la ciencia en 1985. La influencia que Eddington ejerció sobre él queda patente en sus palabras: "En 1923, cuando era estudiante de primer año en Cambridge, un amigo se refirió a Eddington como la persona que más sabía sobre relatividad aparte del mismo Einstein. En 1926 el astrofísico E. A. Milne me aconsejó que, cualquiera que fuera la rama de la física a la que pensara dedicarme, leyera el libro de Eddington *Internal Constitution of the Stars*. Aunque durante algún tiempo trabajé en física matemática, la lectura del libro influyó en mi decisión de dedicarme a la astrofísica y me ha llevado, indirectamente, a escribir este artículo".

Trató el tema libre y profundamente en un buen número de sus conferencias y discursos, donde hacía especial énfasis en la idea de "búsqueda" que caracteriza a la tradición cuáquera.

Nunca sugirió que su estudio de la física y el universo condicionara su perspectiva religiosa. Aceptaba, por supuesto, que la verdad era una y que el hombre busca descubrir la conexión que liga todas las cosas, pero también afirmaba que los científicos quizá tuvieran que recorrer un largo camino antes de encontrar las conexiones esenciales. Como él mismo dijo durante una conferencia en 1929: "En ciencia, como en religión, la verdad resplandece como un faro mostrándonos el camino; no pretendamos alcanzarla, es mucho mejor que se nos permita ir en su busca."

Coherente con su fe, se hizo objetor de conciencia al estallar la primera guerra mundial. Enterado de ello Sir Frank Dyson, Astrónomo Real, confió en poder destinarlo a la observación del eclipse solar de 1919, si la guerra lo permitía. Con este fin, utilizó su influencia para mantener a Eddington en Cambridge durante el conflicto. Otros científicos sobresalientes se negaron, asimismo, al servicio de armas, pero la mayoría cambiaría de opinión al declararse la segunda guerra mundial.



1. EDDINGTON estaba en el momento cumbre de su carrera a principios de los años veinte, cuando posó para este retrato oficial de la presidencia de la Regia Sociedad Astronómica de Londres.

No parece que tuviera una meta científica definida hasta que ingresó en el Observatorio Real en 1906. Se sumó con entusiasmo al trabajo que allí se realizaba, iniciando, por ejemplo, el programa para determinar la variación de la latitud, continuado, de una u otra manera, casi hasta nuestros días.

Dedicó buena parte de los 10 años siguientes al análisis de los movimientos estelares, estudio comenzado en 1904 por el astrónomo neerlandés J. C. Kapteyn. Hizo una revisión completa del tema, incluyendo sus propias contribuciones en la materia, en su primer libro, *Stellar Movements and the Structure of the Universe*, publicado en 1914. Así se imaginaba entonces el universo: en primer lugar estaba el "sistema estelar" o "galaxia", concebido como un conjunto aplanado de estrellas, en el que el Sol ocupaba una posición "más o menos central". El astro estaba rodeado por la Vía Láctea, formada por estrellas, nebulosas y nubes de materia oscura.

El sistema en su conjunto parecía estar en reposo, lo que llevó a Einstein a plantear un modelo estático del cosmos introduciendo la ya famosa constante cosmológica en la versión original de la teoría general de la relatividad. Eddington basaría en esta constante, a la que Einstein renunciaría más tarde, gran parte de su trabajo posterior.

La parte principal del libro, sobre cinemática estelar, proporcionó el punto de partida para un buen número de investigaciones posteriores. El último capítulo del mismo fue el estímulo para otro gran trabajo en dinámica estelar. Probablemente, fue también Eddington el primero que dio una explicación correctamente fundada acerca de la naturaleza de las entonces denominadas nebulosas espirales. Apoyó —correctamente, por supuesto— la hipótesis de que se trataba de galaxias como la nuestra. Llamó la atención sobre la necesidad de suponer, como Henri Poincaré había sugerido, que nuestra galaxia está en rotación. En definitiva, sistematizó todo el conocimiento de la época acerca de la macroestructura del universo conocido.

Tuvo su primer contacto con la relatividad general en 1912, tres años antes de que Einstein publicara una versión completa de su teoría. Había éste obtenido algunos resultados provisionales, a partir de los que dedujo que los rayos de luz provenientes de una estrella, al pasar por las cercanías del Sol, serían desviados por su campo gravitacional. Tal desviación debería ser observable durante un eclipse total de Sol, cuando el disco solar queda oscurecido, siempre que la estrella en cuestión brille lo suficiente para ser vista a través de la corona solar. El desplazamiento que Einstein había calculado en este punto de sus investigaciones era, en realidad, sólo la mitad del valor correcto predicho por la teoría completa y podía haberse obtenido fácilmente mediante una pequeña modificación en la teoría de la gravitación de Newton.

En 1912, E. F. Freundlich, que an-

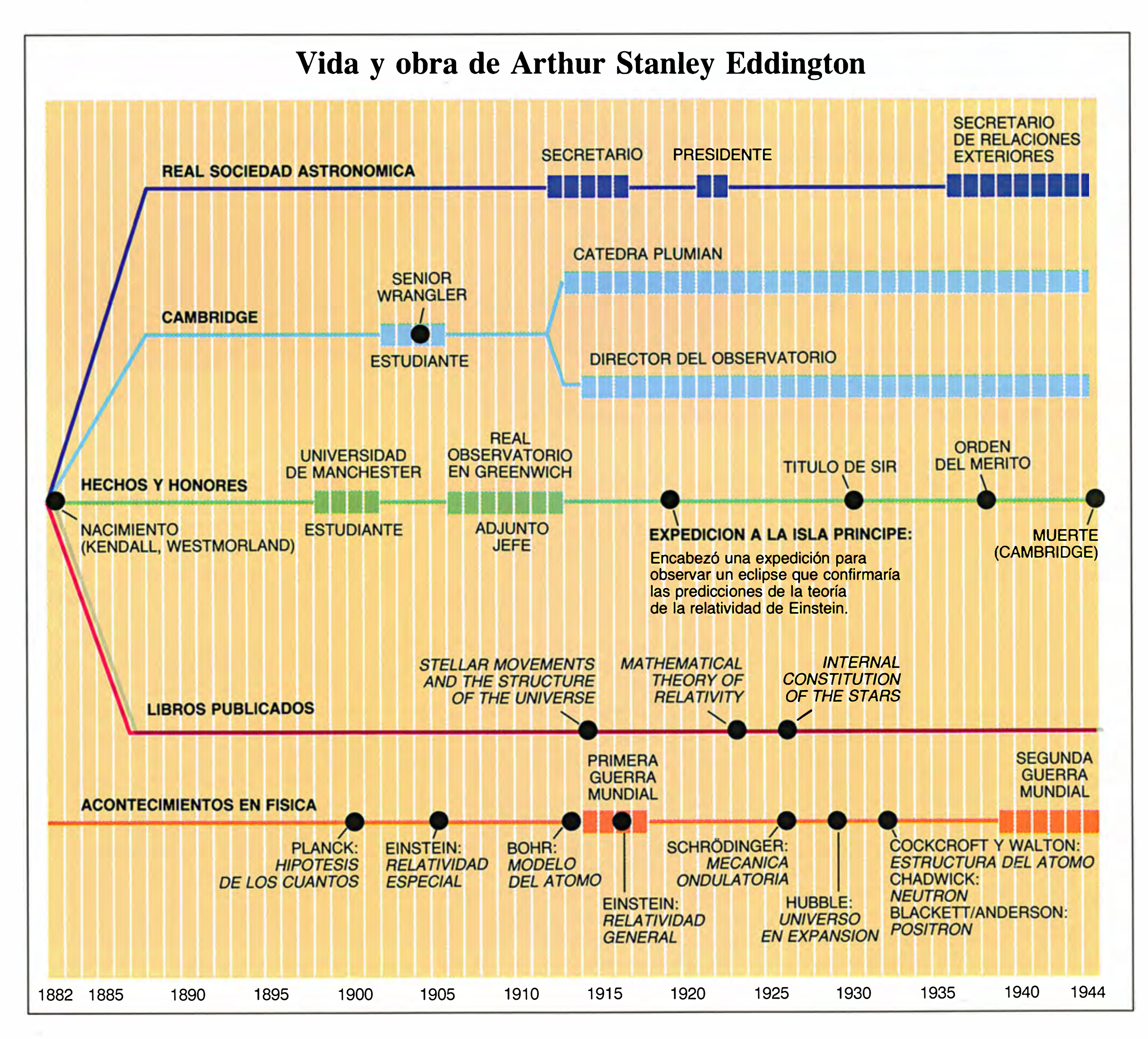
dando el tiempo sería compañero de Einstein, se dispuso a preparar un experimento para comprobar la predicción de Einstein durante un eclipse que sería observable en Europa del Este en 1914. Solicitó asistencia técnica al comité conjunto de eclipses formado entre la Real Sociedad Londinense y la Real Sociedad Astronómica, del que Eddington era miembro y Dyson presidente. El comité lamentó tener que comunicar a Freundlich que no podía disponer del equipo especial que éste necesitaba. Sin embargo, Dyson y Eddington fueron puestos sobre aviso del proyecto que estaba en marcha.

La guerra, que cortó bruscamente el normal intercambio de información científica entre Alemania e Inglaterra, estalló antes de que Freundlich pudiera observar el eclipse. Así, cuando en 1915 se publicó el artículo definitivo de Einstein sobre la teoría de la relatividad general, sólo una copia del mismo llegó a Inglaterra. Se la envió a Eddington, secretario de la Real Sociedad Astronómica en esas fechas, W. de Sitter, eminente astrónomo holandés, quien le remitía, además, tres importantes artículos propios en los que interpretaba los últimos descubrimientos de Einstein y añadía algunas contribuciones propias. Estos artículos fueron los primeros sobre relatividad general publicados fuera de Alemania.

El siguiente lo redactó Eddington: *Report on the Relativity Theory of Gravitation*, presentado en la Sociedad de Física de Londres en 1918. Se nos revela en él la asombrosa capacidad de Eddington para asimilar y explicar conceptos y técnicas matemáticas nuevos y revolucionarios. El trabajo sirvió de base para su *Mathematical Theory of Relativity*, un clásico publicado en 1923. En él han bebido la mayoría de los físicos.

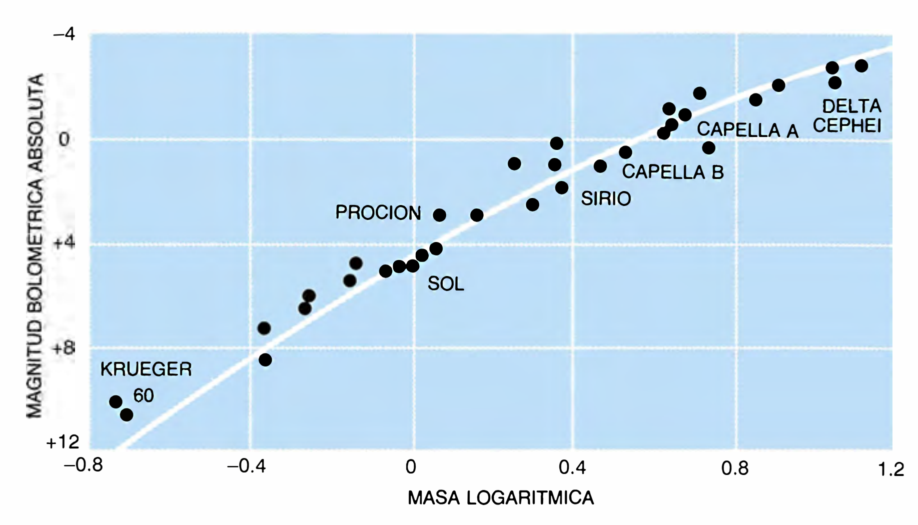
Dyson destacó que el eclipse solar de mayo de 1919 tendría lugar contra el fondo estelar probablemente más favorable, una zona muy rica en estrellas del cúmulo de las Híades. Para su fortuna, el fin de la guerra permitió a Dyson organizar dos expediciones británicas para realizar las observaciones. Eddington encabezó una de ellas a la isla Príncipe, frente a la costa oeste de Africa, que fue la primera en comunicar sus resultados. La otra expedición se dirigió hacia Sobral, en la parte noreste de Brasil. En conjunto, las mediciones resultaron compatibles con los cálculos de Einstein y definitivamente diferentes de lo predicho por la teoría de Newton.

El efecto registrado consistía en un desplazamiento de unas centésimas de milímetro sobre las placas fotográficas. Algunos físicos sostienen que, con la técnica empleada por Eddington, los márgenes de error resultaban demasiado amplios como para confir-



mar una teoría y refutar la otra. Esta crítica es injusta: una confirmación no es una prueba, pero es todo lo que la física experimental puede ofrecer. Para el refinamiento de las técnicas hubo que aguardar al desarrollo de la radioastronomía, tras la segunda guerra mundial.

La necesidad de una modificación de la ley de gravitación universal de Newton produjo una conmoción mundial, posiblemente mayor que cualquier otro descubrimiento anterior o posterior. El sentimiento internacionalista tras el final de la guerra contribuyó a la gran acogida que tuvo la nueva teoría: los británicos apoyaban el trabajo de un científico nacido en Alemania. Einstein saltó inmediatamente a la fama. Y si la ley de Newton debía modificarse, justo era que la causa fuera el trabajo de un grupo de astrónomos británicos.



2. LA LEY MASA-LUMINOSIDAD, obtenida por Eddington, relaciona el brillo de las estrellas de la secuencia principal (como el Sol, en el centro) con sus masas (sólo existe una leve dependencia del radio estelar). Eddington realizó esta predicción sin saber cuál era la fuente de energía de las estrellas.

Diríase, pues, que Eddington fue la primera persona capaz de "asomarse" a las profundidades de una estrella. Hay datos geológicos de que la estrella más cercana a nosotros, el Sol, ha estado radiando energía a un ritmo esencialmente constante durante miles de millones de años. ¿Qué procesos tienen lugar en el interior del Sol para posibilitar este comportamiento? Los contemporáneos de Eddington solían esgrimir el siguiente argumento: "No puedes saber cómo es el interior del Sol si no sabes de dónde obtiene su energía, pero no puedes saber de dónde obtiene su energía a menos que sepas cómo es en su interior." Hasta el simple dato de que la densidad media del Sol era sólo vez y media la del agua resultaba desconcertante. ¿Qué estado de la materia permitiría que un cuerpo dotado de tanta masa poseyera esa escasa densidad? ¿Cómo se transmitía en un medio tal la energía? La hipótesis más aceptada suponía que el transporte se realizaba mediante corrientes de convección.

Eddington comenzó su trabajo sobre estructura estelar en 1916. Quería averiguar por qué las "Cefeidas variables", un grupo de estrellas, producían variaciones periódicas de luminosidad. Modelizó las Cefeidas en esferas de gas que se contraían y expandían alternativamente. Siguiendo una idea expuesta por R. A. Sampson en 1894, consideró que el transporte de energía podría ser radiativo. A partir de aquí desarrolló cierto trabajo, convertido ya en clásico. Más aún: el éxito evidente del modelo le llevó a pensar que la hipótesis de la esfera gaseosa podría aplicarse a otros propósitos más generales.

Así, Eddington se impuso el problema de calcular el estado de equilibrio de un cuerpo autogravitante, de masa y composición dadas, con propiedades de gas ideal y transporte de energía por radiación. Postuló una fórmula para la opacidad del material, en función de su composición y estado físico, y postuló asimismo una distribución inicial de fuentes de energía en su interior. Para entonces, el físico suizo J. R. Emden ya había calculado el estado de equilibrio de un grupo particular de esferas de gas conocidas como politropos. Eddington se dio cuenta de que, bajo ciertas simplificaciones razonables, sus ecuaciones se reducían a las de Emden y, por tanto, podía utilizar sus resultados numéricos. De esta forma obtuvo lo que denominó el modelo estelar estándar.

El modelo de Eddington proporcionaba una relación simple entre la masa, el radio y la luminosidad de una estrella, en la que la dependencia del radio estelar era muy leve. Comprobó entonces que la ley masa-luminosidad resultante se ajustaba a las observaciones de todas las estrellas normales, incluido el Sol, para el que sí poseía datos numéricos precisos. Demostró, pues, que todas las estrellas normales se comportan, esencialmente, como un gas perfecto.

La idea de que un material mucho más denso que los líquidos y sólidos ordinarios, como el que se suponía debía existir en las regiones centrales del Sol, obedeciera la misma ley que el aire que respiramos constituía una noción revolucionaria. Lo descubrió Eddington. Un amigo suyo, J. H. Jeans, llamó la atención sobre la necesidad de que el material que forma las estrellas se hallara altamente ionizado. Eddington mostró que la ionización era el proceso que hacía que esta materia, a muy alta densidad, se comportara como un gas ideal. Podía afirmarse entonces que una cantidad de material estelar con masa igual a la del Sol formaría una estrella con su misma luminosidad, prediciendo así el flujo de energía sin conocer su origen.

Para ello, argumentaba que la energía radiativa debía abrirse camino venciendo la opacidad del material que atraviesa. Esta opacidad depende de su temperatura y densidad según una fórmula fácilmente deducible de la mecánica cuántica. Postuló entonces que el ritmo de generación de energía dependía igualmente de la temperatura y densidad del material. Sus cálculos mostraban que ambas magnitudes debían ajustarse mutuamente para dar lugar a la relación expresada en su ley masa-luminosidad. De no ocurrir tal, la estrella no sería capaz de mantener un estado de equilibrio y su existencia sería imposible.

El modelo de Eddington también implicaba que la energía generada en las zonas más profundas de una estrella como el Sol tardaría aproximadamente unos 10 millones de años en llegar a la superficie de la estrella y escapar. De esta forma, si la generación de energía se interrumpiera bruscamente en el Sol, tardaríamos aproximadamente 10 millones de años en notar que algo había sucedido. Este lapso de tiempo entre la causa y sus efectos visibles es lo que permitió a Eddington deducir las propiedades del interior de una estrella sin conocer su fuente de energía.

Cuando presentó sus conclusiones en la Regia Sociedad Astronómica provocó una aguda polémica. Algunos se sumaron a la crítica de Jeans, quien defendía que la luz emitida por una estrella dependía por entero de las fuentes de energía que la originan. Por tanto, sin conocer una no se pue-

de predecir la otra. Hubo incluso quien destacó lo absurdo de suponer que un material tan denso como un ladrillo pudiera ser tratado como un gas. Un poco más tarde E. A. Milne declaró que algunas de las conclusiones de Eddington derivaban de las restricciones impuestas en el tratamiento matemático, más que de la física del problema.

Los debates desencadenados en las reuniones mensuales de la Regia Sociedad Astronómica a mediados de los años veinte despertaron un inusitado interés. A ellos acudía la flor y nata de los físicos y matemático hubo que se afilió a la academia para poder asistir. Eddington y Jeans no se daban cuartel; podían comportarse así, y disfrutar haciéndolo, porque en privado eran grandes amigos. Pero si algún astrofísico joven entraba en la discusión, entonces radicalizaban sus posturas. El resultado final de todos estos debates fue la aceptación de la tesis de Eddington, quien resumió su pensamiento en *The Internal Constitution of the Stars*, un clásico de la astrofísica teórica publicado en 1926.

En realidad, Eddington había tenido más suerte de lo que suponía. Una docena de años más tarde, se aceptó que las reacciones de nucleosíntesis termonuclear constituían la fuente de energía de las estrellas. Estas reacciones son extremadamente sensibles a la temperatura, sensibilidad que le permitió a nuestro autor describir la estructura interna de las estrellas sin conocer el proceso de generación de energía en su interior.

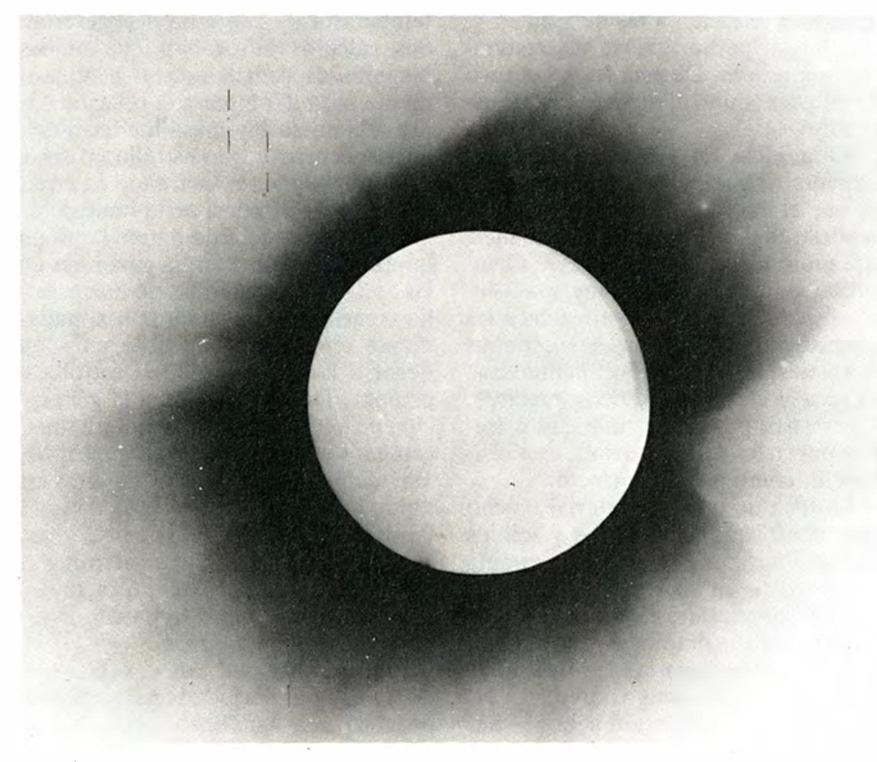
Muchos astrofísicos creen que, hacia 1935, Eddington se equivocó en un punto en particular. El mismo había mostrado cómo su modelo topaba con una dificultad insalvable cuando intentaba aplicarlo a las estrellas más densas, conocidas como enanas blancas (por ejemplo, la compañera oscura de Sirio, la estrella más brillante, después del Sol, vista desde la Tierra). Los físicos matemáticos de Cambridge señalaron inmediatamente que los electrones libres de una estrella tan densa formarían un gas degenerado, en sentido mecánico-cuántico, que resistiría cualquier intento posterior de compresión. El material obedecería entonces una ecuación de estado diferente de la utilizada por Eddington para la materia "clásica". R. H. Fowler mostró que la dificultad quedaba así salvada y Eddington aceptó de buen grado esta explicación.

Otros teóricos, sin embargo, puntualizaron que tan altas presiones elevarían las velocidades efectivas de los electrones degenerados, acercándolas a la de la luz, adquiriendo entonces interés los efectos relativistas. Así, la ecuación de estado para un gas de electrones degenerados y relativistas diferiría de la anteriormente mencionada. Por la misma época, Subrahmanyan Chandrasekhar, entonces en la Universidad de Cambridge, obtuvo un importante resultado teórico: un cuerpo compuesto por este "gas relativista" no puede mantenerse en estado de equilibrio si su masa es mayor de unas 1,5 masas solares ("límite de Chandrasekhar"). A menos que algún proceso consiguiera mantener este cuerpo por debajo de su masa crítica, sobrevendría un colapso catastrófico que daría lugar a la formación de un agujero negro, cuerpo tan denso que su atracción gravitatoria impide que incluso la luz escape de él. No existe ningún dato observacional que entre en conflicto con el resultado de Chandrasekhar.

Eddington no aceptó esta conclusión y la consideró mera reducción al absurdo de la teoría. La mayoría de los astrofísicos piensan que éste fue su mayor error; no pudo encontrar ningún defecto en el trabajo de Chandrasekhar. Sin embargo, en cincuenta años nadie ha calculado detalladamente cómo evolucionaría una porción de gas con masa superior al límite de Chandrasekhar. Tampoco ninguno de los candidatos a agujero negro de los sistemas estelares observados ha sido confirmado como tal. Quizás estamos trabajando justo en los límites de validez de la teoría cuántica y la relatividad estándar. Si estos límites no han sido sobrepasados, indudablemente Eddington estaba equivocado. Pero si no erró, entonces los astrofísicos mencionados habrían obtenido sus conclusiones aplicando la teoría fuera de su rango de validez.

Cuando Eddington comenzó su estudio de la estructura estelar, hace 70 años, los conocimientos sobre la materia eran escasísimos. No deja de sorprender que la visión global que muy pronto llegó siga siendo esencialmente válida en nuestros días. Quizá tuvo un golpe de suerte y quizá se equivocó al evaluar las consecuencias de la degeneración relativista, pero la potencia y alcance de la intuición física de Eddington aún causan asombro.

En su primer gran trabajo astronómico, que culminó con la publicación de *Stellar Movements*, describió la estructura de una galaxia, de cualquier galaxia. Más tarde, en *Internal Constitution* aclaró la composición de las estrellas. Puesto que las galaxias están formadas principalmente por estrellas, podemos decir que toda la astronomía moderna descansa sobre los descubrimientos de



3. EL ECLIPSE SOLAR de 1919 confirmó la predicción de Einstein según la cual la luz de una estrella que pasara cerca del Sol sería desviada por el campo gravitacional de éste. Tal deflexión sólo supone unas centésimas de milímetro en esta fotografía en negativo, tomada en Sobral, Brasil (las estrellas vienen indicadas por las líneas verticales). Eddington encabezó una expedición similar a la isla Príncipe, frente a la costa oeste de Africa.

Eddington. Sin duda, un gran mérito para un solo hombre.

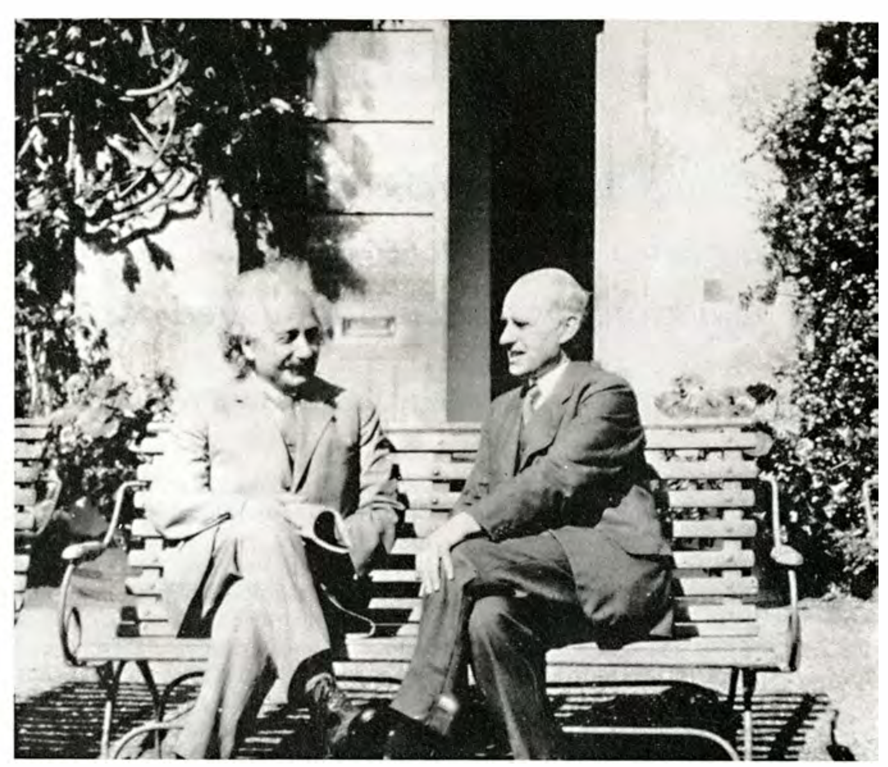
No se detuvo ahí. Demostró que el modelo de universo estático desarrollado por Einstein era inestable, y concluyó que, de acuerdo con la relatividad general, el universo debía estar en expansión (o bien en contracción, pero no estático). Hizo así una contribución esencial a la astronomía del universo en cuanto tal.

Eddington pasó revista a su obra cosmológica instalándose a extramuros de esa disciplina. Comenzó a interesarse por las implicaciones filosóficas de su trabajo, mientras se distanciaba de la labor exclusivamente astrofísica. ¿Qué es el universo? ¿Qué es la física? A tales cuestiones fundamentales dedicó el resto de su vida científica.

Podemos encontrar indicios de sus inquietudes metafísicas ya en 1929, cuando escribió *The Nature of the Physical World*. Pensamiento que se halla más elaborado en *Philosophy of Physical Science*, de 1939.

Se dio cuenta de que algunas constantes físicas no dependían del sistema de unidades en el que estaban expresadas. Citaremos, de tales números "puros", el cociente entre la fuerza de atracción electrostática entre un protón y un electrón y la fuerza de atracción gravitatoria entre ellos (un número del orden de 10^{39}), así como el cociente entre las masas del protón y el electrón (aproximadamente, 1836). Las teorías físicas clásicas consideran que estos cocientes son constantes universales —es decir, tienen un valor preciso— y que podemos aproximar su valor real por medio de mediciones más refinadas, pero nunca obtener su valor exacto.

Eddington puso en cuestión ese punto de vista. Estaba convencido de que deben existir razones matemáticas que justifiquen esos valores de las constantes. Y elaboró una teoría para descubrir estas razones. Los resultados que obtuvo concordaban, a veces increíblemente bien, con las mediciones empíricas. El desarrollo de esta teoría involucraba un tratamiento matemático muy complejo e incluso algunas técnicas que Eddington desarrolló para la ocasión. Pese a ello, nadie pudo encontrar ningún error matemático serio en su trabajo.



4. GRANDES CONTEMPORANEOS: Einstein calificaba los libros de Eddington sobre relatividad como lo mejor que se había escrito en lengua inglesa. En la foto, ambos científicos en el jardín de Eddington, en 1930.

Es cierto que ningún resultado particular de este último trabajo de Eddington ha recibido aceptación. Pero tampoco se sabe exactamente de qué postulado partió para elaborar su teoría. Además, ningún investigador ha podido seguir su línea de razonamiento hasta alguna de sus conclusiones principales; siempre aparecía alguna laguna en la lógica del proceso.

Eddington fue un gran hombre, un gran científico... y un gran oportunista. El científico debe estar alerta a lo que sucede a su alrededor. Algunos de los conocimientos que utilizó parecían haber sido obtenidos para él, por ejemplo, los resultados de Kapteyn sobre el movimiento estelar y de Emden sobre la composición de las estrellas. Desde una perspectiva más general, la física atómica había alcanzado justo el grado de desarrollo necesario para que Eddington pudiera emplearla para comprender la opacidad, la ionización e incluso parte de los mecanismos de generación de energía en una estrella.

Pese a que Eddington tenía intereses comunes con sus contemporáneos de Cambridge, no puede decirse que perteneciera a una escuela de pensamiento en particular. Tampoco inició una línea propia, aunque le sucedieron toda una serie de investigadores formados por él. De sus 110 artículos de investigación, sólo 7 eran colaboraciones.

Incluso en su teoría sobre los fundamentos de la física, Eddington fue uno más del pequeño grupo de investigadores que perseguían, cada uno por su cuenta, los mismos fines. Eddington, Einstein, Milne, Erwin Schrödinger un poco más tarde y Hermann Weyl un poco antes, trataron de construir teorías unificadas para la física y la cosmología sin conseguirlo. Podría pensarse que desperdiciaron su tiempo y su talento en semejante empeño, pero, conociéndoles individualmente, es difícil imaginar qué otra cosa hubieran podido hacer, sino aceptar el desafío. Es más, lo que ellos intentaron es lo que algunos de los mejores vienen haciendo desde entonces: tratar de construir teorías de gran unificación. De hecho, se han obtenido resultados parciales satisfactorios, como la unificación del electromagnetismo y la fuerza débil.

Los pioneros de la primera mitad de este siglo abrieron el camino a estos descubrimientos y, en este aspecto, el más eficiente y profético parece haber sido Eddington.

BIBLIOGRAFIA COMPLEMENTARIA

ARTHUR STANLEY EDDINGTON 1882-1944. H. C. Plummer en *Obituary Notices of Fellows of the Royal Society*, vol. 5, n.º 14, págs. 112-115; noviembre de 1945.

THE DEVELOPMENT AND MEANING OF THE EDDINGTON'S "FUNDAMENTAL THEORY". Noel B. Slater. Cambridge University Press, 1957.

THE LIFE OF ARTHUR STANLEY EDDINGTON. A. Vibert Douglas. Thomas Nelson and Sons, 1957.

MEN OF PHYSICS: SIR ARTHUR EDDINGTON. C. W. Kilmister. Pergamon Press, 1966.

EDDINGTON: THE MOST DISTINGUISHED ASTROPHYSICIST OF HIS TIME. Subrahmanyan Chandrasekhar. Cambridge University Press, 1983.

El cuásar 3C 273

Es uno de los objetos más luminosos del universo conocido y el núcleo de una galaxia activa. Del estudio de su espectro de radiación, los astrónomos infieren por qué brillan los cuásares

Thierry J.-L. Courvoisier y E. Ian Robson

El cuásar 3C 273 está situado aproximadamente a una quinta parte del camino que separa la Tierra del límite del universo conocido. De todos los objetos del cosmos, sólo algunos más, cuásares también, superan la energía y actividad de 3C 273. En un día normal brilla más que 1000 galaxias de 100.000 millones de estrellas cada una. Cierto día del mes de febrero de 1988, irrumpió con un pulso de radiación equivalente a encender estrellas del tamaño de nuestro Sol a un ritmo de 10 millones por segundo.

La observación atenta de 3C 273 en todas las regiones del espectro electromagnético y el registro cuidadoso de las variaciones de su luminosidad han posibilitado que los astrónomos comenzaran a comprender los cuásares y los procesos físicos que les suministran energía.

Desde su identificación hace 28 años, se ha terminado por saber que se trata de los núcleos de galaxias extremadamente activas. Los cuásares no tienen par en luminosidad y, por tanto, son los objetos observables más distantes del universo. De los descubrimientos que les conciernen, podemos destacar uno relativo a su luminosidad: varía muchísimo en períodos inferiores a un año, variabilidad que indujo a concluir que la ingente energía de los cuásares es radiada por una zona mucho menor que los núcleos de las galaxias normales.

Tal energía proviene de la energía gravitacional que se libera cuando el gas y el polvo caen hacia sus centros densos y masivos. Parte de ella canaliza las partículas en haces, lanzando materia hacia la galaxia anfitriona a velocidades próximas a la de la luz. Buena parte de esa energía se convierte en radiación a través de una gavilla de procesos físicos que ocurren, probablemente, a distinta lejanía del núcleo. Además, los cuásares presentan muchos aspectos que no se pueden explicar y siguen contándose entre los objetos que provocan mayor desconcierto del universo.

Sabemos más de 3C 273 que de ningún otro cuásar. Posee un amplio repertorio de propiedades, no todas las cuales son compartidas por los demás cuásares. Ahora bien, el derroche de actividad de que hace gala 3C 273 ayuda a los astrónomos a desentrañar los fenómenos que acontecen en los cuásares.

La observación de 3C 273 constituye una tarea tan ardua cuan agradecida. Después de viajar por el espacio más de mil millones de años, sólo una pequeña fracción de su radiación alcanza nuestro planeta. Para captarla se requiere un seguimiento tenaz, acometido por una batería de telescopios terrestres e instrumentos instalados a bordo de satélites.

El empeño ha cumplido más de un siglo. El objeto bautizado ahora como 3C 273 conoció su primer registro en placas fotográficas cuando los astrónomos analizaban las estrellas de la constelación de Virgo. Parecía una estrella más, de brillo moderado. En 1962, Cyril Hazard y sus colaboradores, de la Universidad de Sidney, descubrieron que el objeto estelar ocupaba la misma posición en el cielo que una fuente intensa de radioondas. El radioemisor llevaba la etiqueta de 3C 273, es decir, el número 273 en el Tercer Catálogo de Cambridge de radiofuentes. A ese tipo de objetos se les llamó luego "radiofuentes cuasiestelares", o cuásares.

En 1963, Maarten Schmidt, de los observatorios Monte Wilson y Palomar, dedujo que el cuásar 3C 273 quedaba a unos 3000 millones de años-luz de la Tierra. Lo que tal hallazgo implicaba resultaba extraordinario. El cuásar era, con mucho, el objeto más luminoso y distante jamás observado. Pronto se identificaron otros cuásares que parecían más alejados y más brillantes que 3C 273. Por aquella época, los colegas de Schmidt tenían buenas razones para cuestionar sus resultados. Pero, conforme se van revisando los datos recogidos durante la última treintena de años, encontramos poco margen para dudar del acierto de Schmidt.

Los astrónomos están ya convencidos de que los cuásares no constituyen un fenómeno aislado. Se trata de los objetos celestes más activos que hallamos en el núcleo, o centro, de las galaxias. Apreciaremos mejor sus propiedades, si las comparamos con las de las galaxias normales.

Las galaxias, nuestra Vía Láctea entre ellas, abarcan unos 100.000 millones de estrellas. En su mayoría, la radiación observada desde la Tierra proviene principalmente de las estrellas que la forman. Por tanto, la luminosidad galáctica viene a ser la suma de las luminosidades de las estrellas componentes. En un segundo grupo de galaxias, parte de la radiación se origina en los gases interestelares iluminados por estrellas calientes. Y hay otro grupo en las que el núcleo genera la mayor parte de la radiación, que podría incluso superar a todas las estrellas asociadas. De estas galaxias activas, así se llaman, los cuásares son el caso extremo.

El espectro de radiación de las galaxias (en realidad, el de cualquier objeto caliente) presenta tres clases de rasgos: el continuo, las líneas de absorción y las líneas de emisión. La

THIERRY J.-L. COURVOISIER y E. IAN ROBSON han colaborado en muchos proyectos de investigación del cuásar 3C 273 y de otros núcleos de galaxias activas. Courvoisier se halla adscrito al Observatorio de Ginebra. Después de doctorarse por la Universidad de Zurich, en 1980, trabajó para la Agencia Europea del Espacio. Robson dirige el departamento de física y astronomía del Instituto Politécnico de Lancashire. En 1973 se doctoró en el Queen Mary College; allí permaneció otros cinco años, hasta que se trasladó al Politécnico de Lancashire. En 1980 fue nombrado director de sus observatorios.

radiación continua está formada por fotones de todas las longitudes de onda. Su intensidad varía suavemente de las longitudes de onda largas a las cortas. En la longitud de onda de una línea de absorción, la intensidad de la radiación es significativamente menor que la de la radiación continua asociada; se originan cuando los gases interpuestos absorben la emisión continua a una longitud de onda determinada. En la longitud de onda de una línea de emisión, la intensidad de la radiación es significativamente mayor que la emisión continua, y se generan, por ejemplo, cuando los gases interestelares absorben emisión continua para reemitirla luego en forma de radiación de cierta longitud de onda.

El espectro de una galaxia normal está formado principalmente por emisión continua, cuya intensidad máxima se produce en el dominio del visible. Por contra, la emisión continua de las galaxias activas y los cuásares es muy intensa desde el infrarrojo hasta la zona de los rayos X y, a menudo, es intensa a las longitudes de onda de radio. Los espectros de las galaxias normales suelen presentar líneas de absorción, pero no de emisión. Sin embargo, las galaxias activas y los cuásares exhiben intensas líneas de emisión en las zonas visible y ultravioleta.

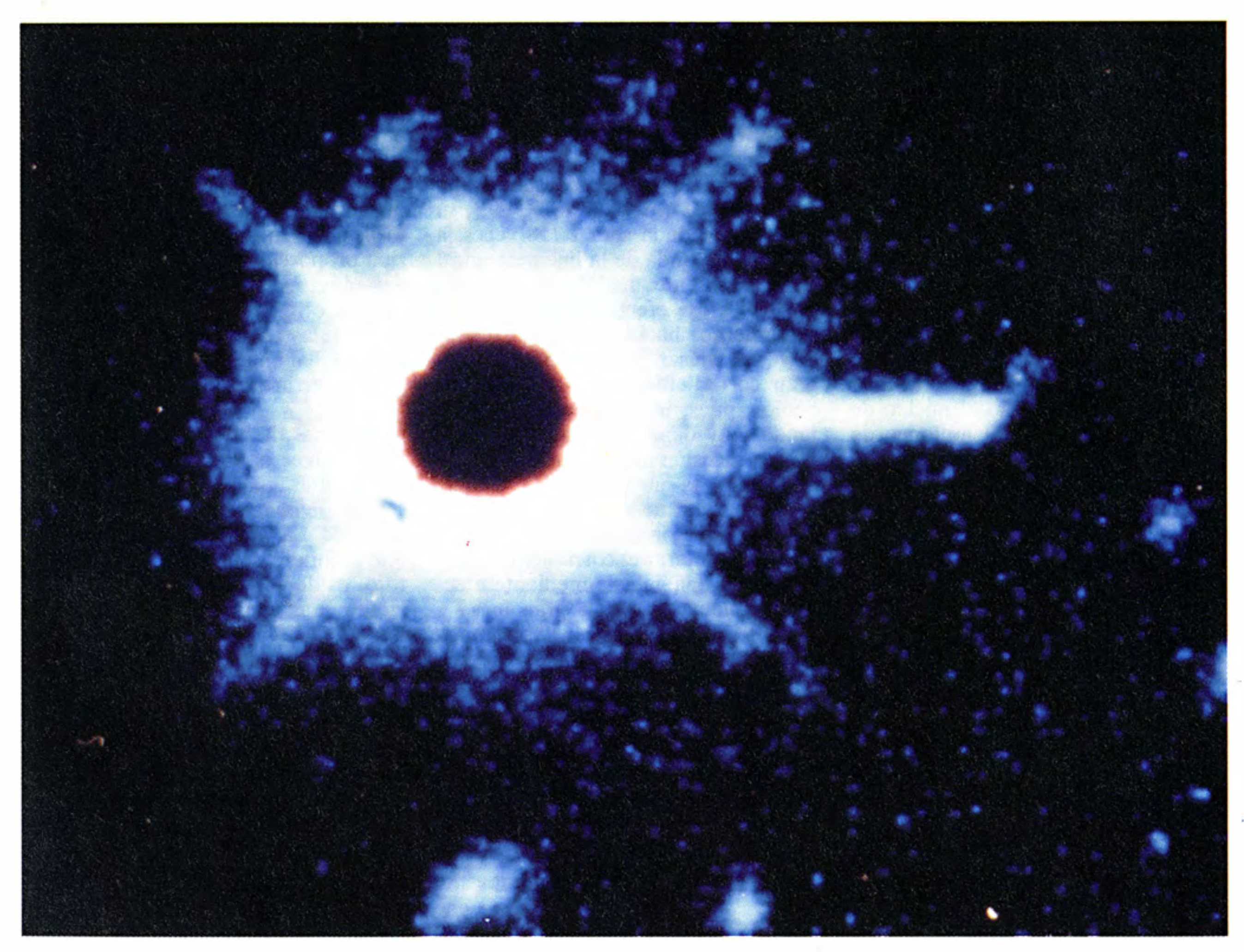
En las galaxias activas, las líneas de emisión se originan cuando la emisión continua ultravioleta y de rayos X ilumina intensamente las nubes gaseosas. Los átomos del gas absorben la energía de la luz ultravioleta y de los rayos X y, después, emiten radiación a longitudes de onda bien definidas. Por tanto, las líneas de emisión indican que la galaxia activa posee una intensa fuente continua de ultravioleta y de rayos X cercana a nubes de gas.

El mismo centro de las galaxias normales puede presentar algunas de las propiedades de las más activas, aunque a escala bastante menor. Es posible que los fenómenos observados en los cuásares ocurran de manera mucho más débil en el centro de nuestra propia galaxia. Cabe también que muchas galaxias normales desarrollaran mayor actividad en el pasado y que la débil ahora observable sea reliquia de un cuásar muerto o adormecido.

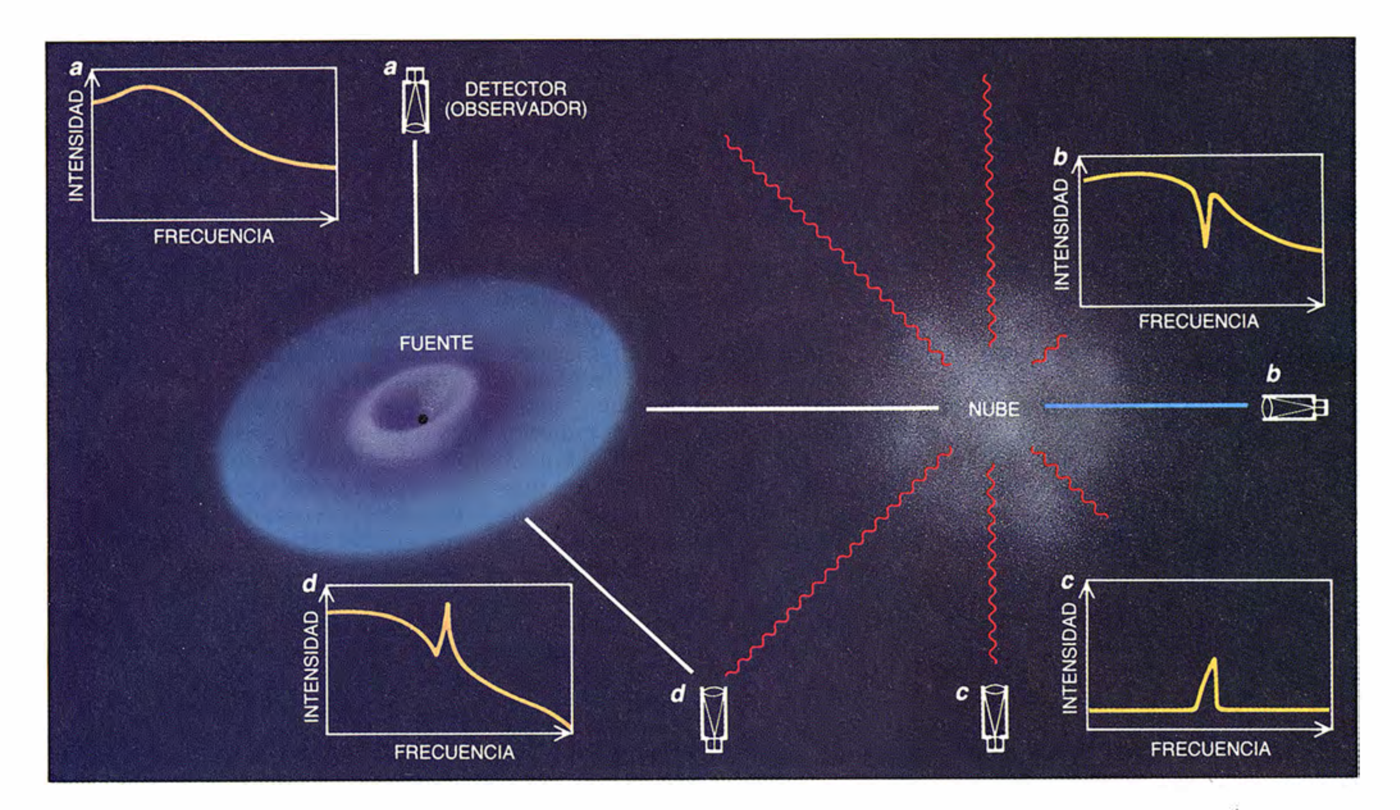
Del estudio de las líneas de emisión de los cuásares y las galaxias activas, podemos sacar información valiosa sobre el tamaño, la estructura y la dinámica de los cuásares.

Algunas de las líneas de emisión observadas en los espectros de los cuásares son anchas, en contraposición a los picos agudos generados por los gases en condiciones de laboratorio. Este ensanchamiento implica que las nubes gaseosas del cuásar se agitan con violencia; su medición permite deducir que las nubes están viajando a miles de kilómetros por segundo. Lo más probable es que esas altas velocidades impliquen la existencia de una fuerza gravitacional intensa causada por un objeto dotado de una gran masa.

La velocidad de los gases nos indica la masa asociada a la fuente central,



1. CHORRO DEL CUASAR 3C 273, cuya proyección apunta hacia la parte derecha de la fotografía. (Las otras líneas débiles, que forman una aspa, son espúreas.) Para realzar la imagen del chorro, los astrónomos han ocultado la luz brillante del núcleo.



2. LOS RASGOS ESPECTRALES se manifiestan al contemplar de distinta manera una fuente de radiación continua y una nube de gas. Un observador próximo a la fuente ve un espectro de radiación continua (*a*). El que mira a través de la nube mide un espectro de líneas de absorción (*b*), producido porque la nube absorbe radiación a ciertas longitudes de onda. Esta luz es reemitida y un tercer observador que mire la nube ve un espectro de líneas de emisión (*c*). Un cuarto, que mire a la vez la nube y la fuente, verá la línea de emisión superimpuesta a la radiación continua (*d*).

lo mismo que la velocidad de los planetas nos habla de la masa del Sol. Mas para calcular con garantía la masa de la fuente central hemos de conocer la distancia del gas al centro, que, a su vez, depende de otros parámetros todavía por averiguar. A pesar de todo, los astrónomos pueden estimar la masa del núcleo de 3C 273. Si ciframos la velocidad de una nube en 5000 kilómetros por segundo a una distancia de 10 años-luz del centro, se calcula que la masa del núcleo de 3C 273 multiplica 2000 millones de veces la solar.

Uno de los puntos más debatidos en torno a los cuásares era el relativo a su lejanía de la Tierra. La distancia que media entre nuestro planeta y una galaxia guarda relación con la velocidad de alejamiento de la galaxia. La ley de Hubble, así se llama dicha relación, es una consecuencia de la expansión cosmológica del universo. Cuando se observa la luz emanada de galaxias alejadas, se está también mirando hacia el pasado, en virtud del tiempo que tarda la luz en viajar de la galaxia hasta la Tierra. Vemos esas galaxias remotas tal como eran en una época en la que el universo era más joven, más pequeño y se expandía más deprisa que en la actualidad. Por tanto, cuanto más alejada esté la galaxia, más deprisa se apartará de nosotros.

Puesto que las galaxias se alejan de la Tierra, las líneas espectrales que emiten están desplazadas. A modo de ejemplo: si observamos desde la Tierra una línea espectral de 3C 273, su longitud de onda será 1,158 veces la longitud de onda de la línea espectral correspondiente medida en el laboratorio; a ese efecto se le denomina corrimiento hacia el rojo. Ahora bien, el cuásar 3C 273 muestra un gran corrimiento hacia el rojo; podemos, pues, inferir que se está alejando muy deprisa. Una alta velocidad de recesión implica, a su vez, que el cuásar está muy alejado.

Punto débil de este razonamiento, y motivo de controversia, es la posibilidad de alterar el corrimiento hacia el rojo de una galaxia mediante efectos relacionados con la gran masa y energía de su núcleo. Por ello algunos afirmaron que los desplazamientos hacia el rojo debíanse, en buena parte, a efectos gravitacionales y otros fenómenos.

Si tales efectos causaran grandes corrimientos de los cuásares hacia el rojo, el desplazamiento hacia el rojo asociado al núcleo de una galaxia debería ser mayor que el producido por sus bordes. Sin embargo, en todos los casos en que se pueden observar galaxia y núcleo activo, el desplazamiento hacia el rojo del segundo ha resultado coincidir con el de la galaxia entera. Este dato avala la idea según la cual, cuando el núcleo activo o el cuásar brillan tanto que impiden discernir la galaxia correspondiente, la distancia deducida del desplazamiento hacia el rojo de las líneas de emisión proporciona la distancia correcta del objeto. Este razonamiento, sumado a otros, ha convencido a la mayoría de astrónomos de que los cuásares se hallan a la distancia indicada por su corrimiento hacia el rojo.

Así pues, la distancia de 3C 273 se deduce de la ley de Hubble. La distancia es igual a la velocidad de recesión dividida por la constante de Hubble. La velocidad de recesión, que se infiere del corrimiento hacia el rojo, es de 44.700 kilómetros por segundo. Las estimaciones de la constante de Hubble están comprendidas entre 15 y 30 kilómetros por segundo por millón de años luz. Por tanto, la distancia de 3C 273 está comprendida entre 1500 y 3000 millones de años-luz.

En otro cálculo, que toma en consideración la distancia entre 3C 273 y la Tierra y el flujo de radiación observado, se descubre que la luminosidad del cuásar multiplica 10^{14} veces la solar: unas 1000 veces mayor que el brillo de una galaxia normal.

Aunque débiles en la zona de radio, los cuásares emiten en casi todos

los dominios del espectro electromagnético: radio, infrarrojo, visible, ultravioleta, X y gamma. Nuestro 3C 273 emite casi la misma energía en todas las regiones comprendidas entre el régimen de radio y el gamma.

El flujo de radiación de 3C 273 y de los demás cuásares puede variar drásticamente en períodos inferiores a un año. Puede cambiar también a velocidades diferentes en zonas distintas del espectro. Esta variabilidad, uno de los primeros rasgos observados de los cuásares, constituye un elemento clave para entender su comportamiento.

Podemos desdoblar el espectro de un cuásar en varias componentes de emisión, cada una de las cuales representa una parte cuya intensidad cambia con idéntica velocidad. Es probable que cada componente provenga de un proceso físico distinto del cuásar. Gracias a la observación tenaz de 3C 273 durante varios años, hemos logrado identificar muchos componentes de emisión distintos. Algunos ya se conocían, mientras que otros se obtuvieron como resultado de estos estudios.

El cuásar 3C 273 se deja observar en el dominio visible desde diciembre hasta julio, momento en el que se acerca tanto al Sol en la esfera celeste que nuestra estrella enmascara su luz. Sin embargo, se pueden detectar radiaciones de longitud de onda mayor durante un período más dilatado.

Con un radiotelescopio instalado en Finlandia, registramos la componente de máxima longitud de onda, que aparece en una serie de picos de la zona de radio. Estas características se deben, así se cree, a nubes de electrones que atraviesan un campo magnético intenso a velocidades cercanas a la de la luz. El campo magnético actúa sobre los electrones y les obliga a seguir trayectorias curvadas y emitir radiación. Los astrónomos atribuyen al proceso de sincrotrón, nombre de ese fenómeno, las emisiones en las zonas de radio y milimétricas. Para observar 3C 273 a longitudes de onda milimétricas, utilizamos el telescopio James Clerk Maxwell de Hawai y el telescopio submilimétrico sueco del Observatorio Europeo del Sur (ESO) de La Silla, Chile.

En 1986, estudiamos la variabilidad del flujo infrarrojo con el telescopio infrarrojo del Reino Unido en Hawai y los instrumentos del ESO en La Silla. Descubrimos que la emisión de sincrotrón no podía contribuir de modo apreciable al flujo registrado en el infrarrojo cercano; ese hallazgo concordaba con observaciones previas, muy desconcertantes, según las cuales la emisión infrarroja no estaba polarizada. La emisión de sincrotrón está, por contra, fuertemente polarizada.

No andamos todavía seguros sobre el origen de la componente infrarroja. ¿Proviene del polvo cósmico calentado a una temperatura de unos 1500 kelvin? Si esta hipótesis fuese correcta, la región polvorienta será muy extensa. El flujo infrarrojo cercano no debería, entonces, variar de forma importante con el tiempo, predicción que nuestras observaciones actuales corroboran.

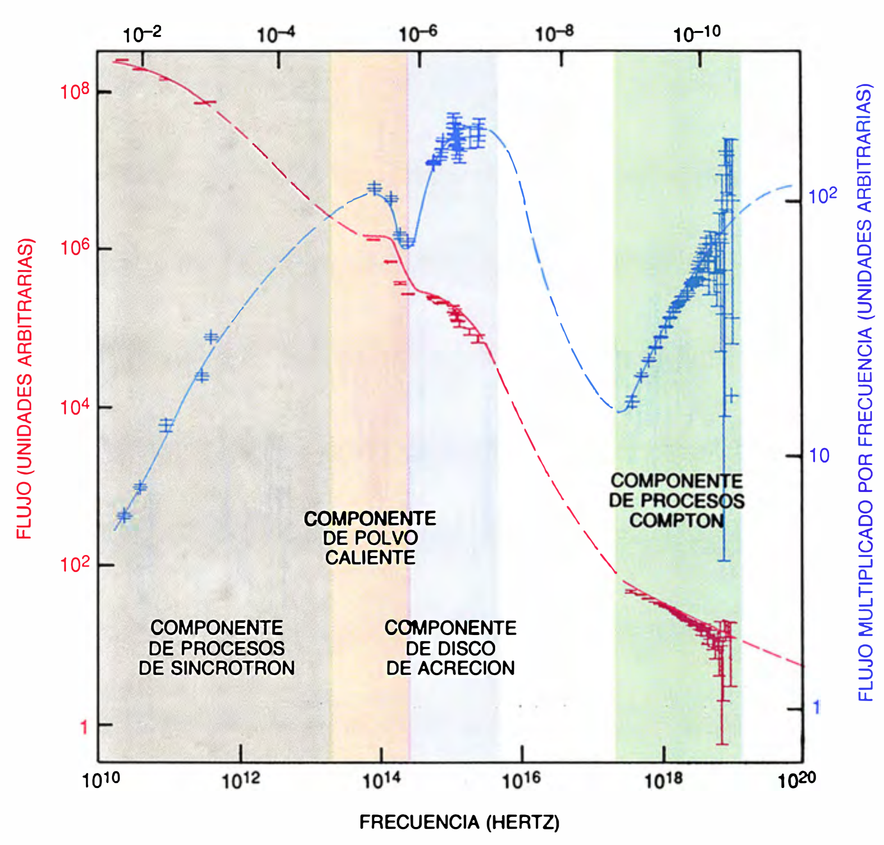
Con el empleo combinado del telescopio óptico suizo de La Silla, un telescopio óptico de las islas Canarias y el satélite *Explorador Internacional Ultravioleta*, hemos seguido la evolución del gran pico azul, una componente intensa y plana. Su origen probable es la emisión superficial de un objeto muy caliente.

Los astrónomos se hallan incapacitados para obtener el espectro de 3C 273 comprendido entre las zonas ultravioleta y de rayos X. Por dos razones: primera, los instrumentos que captan radiación en el dominio ultravioleta extremo no suelen ser tan sensibles como los telescopios que analizan las ondas de radio o las luminosas, y segunda, la radiación de esta banda se absorbe muy fácilmente por la materia interestelar, de manera que el flujo que llega a la Tierra es pequeño en comparación con el recibido a otras longitudes de onda. Un nuevo telescopio espacial, el *Satélite de Observación Roentgen* (*ROSAT*), ha comenzado a operar en ese dominio del espectro.

En las zonas de rayos X y gamma, el flujo decrece según una razón potencial al decrecer la longitud de onda. La componente de rayos X está producida, probablemente, por emisión Compton inversa, fenómeno en el que intervienen los electrones de alta energía del interior del cuásar. Tales electrones chocan con fotones, a los que ceden parte de su propia energía. En el proceso, los fotones de baja energía se transforman en rayos X.

El *Satélite Europeo de Observación de Rayos X* (*EXOSAT*) nos ha permitido echar un vistazo a la mayor parte de la zona de los rayos X, aunque obstáculos técnicos hacen difícil controlar las longitudes de onda cortas, o rayos X duros. Se han realizado otros registros en ese dominio con el satélite japonés *Ginga*.

Para establecer en qué cuantía cada



3. FLUJO DE RADIACION que llega a la Tierra desde el cuásar 3C 273, medido a distintas longitudes de onda (*línea roja*). Este espectro de energía se puede dividir en cuatro componentes, cada una de las cuales proviene de un proceso físico diferente. Al multiplicar cada valor del flujo por su frecuencia, se obtiene la línea azul, cuyos picos representan regiones donde se emite la potencia máxima. Todas las observaciones se realizaron en julio de 1987, excepto las de la zona de longitud de onda milimétrica. Las líneas verticales muestran el intervalo de incertidumbre de cada medida.

componente de emisión contribuye a la luminosidad total del cuásar, basta un sencillo cálculo: la luminosidad a una longitud de onda determinada guarda relación con la longitud de onda multiplicada por la densidad de flujo a dicha longitud de onda, medida desde la Tierra. Tales cálculos muestran que todas las componentes de emisión tienen luminosidades comparables; fenómeno sorprendente, pues parece que cada componente se origina en un proceso físico diferente. Pero hay dos componentes que dominan: el ultravioleta extremo y los rayos X duros. Por un capricho del destino, estas dos regiones son las más difíciles de observar, lo que nos cela saber la energía total emitida por los cuásares.

La escala temporal de variación de cualquier objeto constituye un dato decisivo, ya que puede indicar el tamaño aproximado de la fuente de la radiación. Como ejemplo sencillo, consideremos una hilera de diez bombillas. Si queremos una caída importante de la luminosidad total, apagaremos muchas: unas seis. Bastará con dar una señal. En el mundo físico, ninguna señal puede viajar más deprisa que la velocidad de la luz. El proceso de apagado durará, al menos, el tiempo que la luz necesita para recorrer la distancia desde el centro de la hilera de bombillas hasta la bombilla más alejada que se tenga que apagar. De la misma manera, el proceso de apagado de los objetos astronómicos durará, al menos, el tiempo que la luz tarda en cruzar la zona de emisión. Por tanto, puede obtenerse un límite superior del tamaño de la fuente multiplicando la velocidad de la luz por la escala temporal de la variación.

Las fotografías de 3C 273 obtenidas a lo largo de los últimos cien años indican que el flujo visible del objeto varía con muchas escalas temporales distintas; la mayoría, por encima de 10 días. También hemos observado recientemente que la mayor parte de las regiones espectrales varían en escalas temporales comprendidas entre un mes y algunos meses. ¿Qué nos dice esa variación? Que el diámetro de la fuente de emisión continua será mucho menor que un año-luz y menor que la distancia del sistema solar a la estrella más cercana.

Nos sorprendió descubrir que las variaciones de las componentes de emisión no mostrasen correlación. Revelábase con ello que la estructura de 3C 273 encerraba mayor complejidad que la esperada. Complejidad que constituye una de las razones por las que los astrónomos han encontrado muy difícil dilucidar las relaciones entre la geometría de los componentes de emisión y su origen físico.

A propósito de esas variaciones la campaña de observación de 1988 nos regaló un hallazgo insólito. En enero observamos que 3C 273 atravesaba un estado de baja actividad. De repente, en febrero, se encendió brusca y reiteradamente. El flujo visible e infrarojo varió hasta en un 50 por ciento. La escala temporal de estos cambios no tenía precedentes: las variaciones más rápidas se producían en el transcurso de un día. No cabía esperar semejante comportamiento en un objeto del brillo de 3C 273 (aunque ya se habían visto antes destellos rápidos en la zona de rayos X antes de algunos núcleos de galaxias de baja luminosidad).

La violenta actividad duró cuatro meses; en los mismos se presentaron cinco picos en el espectro óptico. El intervalo entre máximos de flujo se cifraba, en promedio, en 15 días. Dos de los máximos sólo estaban separados, sin embargo, por dos días. La variación más rápida del flujo fue una disminución de un 15 por ciento en 24 horas. ¡Equivalía a apagar unos 10 millones de soles por segundo!

Se registraron llamaradas repetidas en la zona infrarroja, pero sólo se detectaron dos máximos porque la componente infrarroja se observaba con menos frecuencia que la región óptica. Ambos coinciden con dos de los máximos del visible. Un máximo infrarrojo venía a duplicar la emisión "en calma". El cambio más rápido aquí observado fue un aumento de aproximadamente el 40 por ciento en 24 horas. Este ritmo de cambio corresponde al encendido de unos 10 millones de soles por segundo.

Este período tan activo aconteció en el momento más oportuno: cuando 3C 273 estaba lejos del Sol y cuando nuestro grupo estaba preparado para reaccionar con presteza. Logramos cubrir el suceso completamente, algunas veces haciendo observaciones diarias. A pesar de todo, el chisporroteo fue tan rápido que sospechamos que muchas de sus características se nos escaparon.

En mayo, 3C 273 volvió a su estado

4. FLUJO observado del cuásar 3C 273 en la zona visible. Varía mucho con el tiempo. En febrero de 1988, cambió rápida y espectacularmente, prueba de que el cuásar se hallaba en un estado de violenta actividad. Las líneas verticales muestran el intervalo de incertidumbre de cada medición.



normal, para permanecer así durante el resto de la campaña de observación. Estamos analizando cuidadosamente períodos anteriores y posteriores a los máximos de 1988 para averiguar por qué 3C 273 entró de repente en actividad. Las observaciones de 1988 fueron los primeros registros disponibles de una actividad tan violenta en los cuásares, prueba de que se trata de un estado relativamente raro o de que los astrónomos no habían tenido hasta entonces la suerte de captar uno de ellos. Con el fin de determinar su grado de rareza, precisaremos en adelante controlar 3C 273 y otros cuásares durante muchos años.

La variabilidad rápida suele manifestar que las regiones de la fuente se mueven a velocidades cercanas a la de la luz (velocidades relativistas). Antes de comprender el fenómeno, hay que saber algo acerca del comportamiento de los electrones de alta energía y del concepto de temperatura de brillo.

Los electrones de alta energía de una fuente muy compacta se enfrían conforme van chocando con los fotones infrarrojos producidos por la emisión de sincrotrón. Los electrones ceden parte de su energía a los fotones y los desplazan a la zona de rayos X. Este proceso combinado, conocido como emisión de sincrotrón auto-Compton, está relacionado con el tamaño de la fuente. En concreto, a medida que el tamaño de la fuente se contrae, aumenta la razón entre la componente auto-Compton y la componente sincrotrón.

La luminosidad y el tamaño de una región de emisión pueden expresarse mediante la temperatura de brillo. No se trata de la temperatura real del objeto sino, sencillamente, de una manera conveniente de expresar la luminosidad y el tamaño. Un objeto que presente emisión de sincrotrón auto-Compton no superará en mucho, o durante mucho tiempo, una temperatura de brillo de 10^{12} kelvin. A temperaturas por encima de ésta, la radiación se emite sobre todo mediante emisión auto-Compton, que enfría de nuevo el objeto hasta 10^{12} kelvin. Además, un objeto cuya temperatura de brillo supere los 10^{12} kelvin emitirá la mayor parte de

su flujo en la zona de los rayos X y no en la zona infrarroja. Por tanto, 10^{12} kelvin es básicamente la máxima temperatura de brillo posible, por lo que se la designa técnicamente como el límite de Compton.

Sin embargo, en algunos casos, la temperatura de brillo observada puede superar los 10^{12} kelvin. ¿Cómo lo consigue? La interpretación más sencilla es que la fuente de emisión se esté moviendo a velocidades relativistas hacia el observador y, en consecuencia, la radiación observada se concentre en un haz y resulte muy intensificada en comparación con la radiación de una fuente idéntica en reposo. A este cambio aparente se le llama "efecto haz".

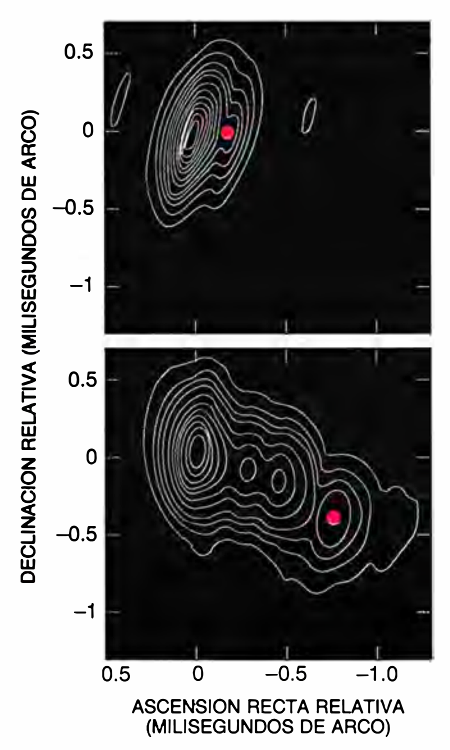
El flujo medido de la fuente es entonces mayor que el que se detectaría si se moviera a velocidades no relativistas. Por tanto, la temperatura de brillo registrada en una fuente relativista sería mayor que la de otra no relativista. De ese modo, la temperatura de brillo observada puede superar el límite de Compton de 10^{12} kelvin.

Las variaciones muy rápidas del flujo óptico e infrarrojo que presenta 3C 273 implican una temperatura de brillo aún muy por debajo del límite de Compton. Sin embargo, las observaciones a longitudes de onda mayores que el infrarrojo sugieren que la temperatura de brillo producida durante los sucesos de febrero y marzo de 1988 estaba muy por encima del límite Compton, por lo que se sospecha que hubo "efecto haz". La observación de que la radiación a longitudes de onda milimétricas varía rápidamente (en una escala temporal de varias horas) proporciona argumentos adicionales e implica temperaturas de brillo altísimas, que superan en mucho el límite Compton. La fuente de la llamarada se encaminaba manifiestamente, a velocidades relativistas, hacia la Tierra.

Hay otro argumento de peso que apunta hacia la existencia de movimientos relativistas en 3C 273. La interferometría de muy larga base (VLBI, de "Very-Long Baseline Interferometry") permite levantar mapas de radio de finísima resolución. La red VLBI ha proporcionado imágenes de radio de varios cuásares; revela, en la mayoría de los casos, núcleos compactos y chorros formados por sucesivas "descargas" de emisión de sincrotrón.

El rasgo más sobresaliente de estas manchas en 3C 273 y otros cuásares yace, diríase, en su alejamiento del núcleo a velocidades varias veces la lumínica. La teoría de la relatividad especial mantiene que las velocidades físicas no superan la velocidad de la luz en ningún lugar del universo. Ahora bien, si un objeto lanza materia en la dirección del observador a velocidades ligeramente menores que la de la luz, se produce la impresión de que la materia avanza a velocidades que superen este límite fundamental.

Acaba de descubrirse que se produjo una nueva descarga, contemporánea a las rápidas variaciones. La red VLBI observó el cuásar 3C 273 en junio de 1988 y en marzo de 1989. La primera dio una imagen de la fuente en la que había aparecido una nueva protuberancia. La imagen de 1989 reveló que se había alejado del núcleo. Con esos datos, los investigadores dedujeron su velocidad aparente y, por tanto, el momento en que fue expulsada del núcleo: la nueva componente emergió durante el período en que 3C 273 estaba lanzando llamaradas de radiación en la zona óptica e infrarroja. De ello podría inferirse que la actividad violenta guarda una estrecha relación con el nacimiento de componentes del chorro. Esperamos que esa conexión nos ayude a entender la formación de los chorros.



5. IMAGENES DE RADIO obtenidas el 25 de junio de 1988 (*arriba*) y el 9 de marzo de 1989 (*abajo*): nos muestran una descarga de radiación (*punto rojo*), que se aleja del núcleo del cuásar 3C 273. Estas imágenes permiten deducir su velocidad y fecha de nacimiento. Parece que se generó durante el violento período del cuásar en febrero de 1988. Las ilustraciones se basaron en datos de T. Krichbaum, del Instituto Max Planck de Radioastronomía de Bonn.

Toda teoría coherente sobre la actividad de los cuásares debe explicar la presencia de chorros relativistas, las líneas de emisión anchas y estrechas y las componentes de emisión continua. Ninguna lo ha conseguido hasta ahora. De acuerdo con la explicación en vigor, la emisión continua de alta energía, radiada por el centro del cuásar, ilumina las nubes de gas que se encuentran a cierta distancia. Los astrónomos pueden explicar, hasta cierto punto, los procesos que originan la emisión continua en términos de emisión de sincrotrón, emisión Compton, etc. Pero nada de eso sirve para conocer dónde se genera la cantidad ingente de energía que necesitan estos procesos. En otras palabras, aunque sabemos que los electrones dotados de mucha energía provocan la radiación, seguimos ignorando cómo la adquieren.

Los investigadores se enfrentan a una labor difícil: justificar que un objeto de un tamaño de sólo unos años luz, e incluso menos, emita la radiación de mil galaxias.

La mayoría opina que los cuásares funcionan, en última instancia, gracias a la energía gravitacional asociada con los agujeros negros, objetos de masa y densidad inmensas. Las pruebas son indirectas. Las rápidas variaciones apuntan hacia una fuente central muy compacta; las estimaciones de masa deducidas de los movimientos de las nubes de gas aportan otras pistas.

Los astrónomos recuerdan también la posibilidad de extraer mucha más energía de las fuerzas gravitacionales que de cualesquiera otras. Han llegado a esta conclusión, en parte, estudiando los procesos gravitacionales y nucleares de los "pulsos" de rayos X de nuestra galaxia. Los núcleos de estos objetos son cuerpos extremadamente densos y calientes: las estrellas de neutrones. A medida que la materia cae hacia la superficie de la estrella, se libera energía gravitacional. Luego, esta materia sufre un proceso de combustión nuclear, que produce pulsos o llamaradas, de rayos X. El análisis de estos procesos pone de manifiesto que las fuerzas gravitacionales liberan cien veces más energía que las nucleares. Sería, pues, natural que la energía gravitacional fuese también la fuente primaria en los cuásares.

La emisión de cualquier objeto compacto está sometida a condiciones. Uno de sus límites es que la propia radiación ejerce una presión. Las fuentes extremadamente luminosas pueden emitir una radiación tan in-

tensa que expulse la materia circundante. Si un objeto gana energía conforme atrae materia hacia él, la luminosidad resultante no podrá superar el límite de Eddington, así se llama el grado de luminosidad con el que la presión de radiación empezará a empujar la materia hacia fuera. Al aumentar la masa del objeto, crecerá también el límite de Eddington. En virtud de ello, concedido que la luminosidad de un objeto está próxima al límite de Eddington, podremos estimar su masa.

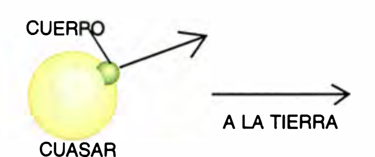
Para el cuásar 3C 273, la masa deducida de la luminosidad multiplica miles de millones de veces la solar. Ahora bien, la masa estimada a partir de la luminosidad de 3C 273 y de otros cuásares no contradice la obtenida en el estudio de movimientos de gases. Dentro de las incertidumbres del caso, resulta razonable pensar, en suma, que la gravitación ejerce un papel importante en la liberación de energía de los cuásares.

Otro parámetro que condiciona la emisión de energía de un cuásar es el ritmo de acreción: la cantidad de masa que cae hacia el centro gravitacional por unidad de tiempo. Los cuásares demandan una gran cantidad de energía gravitacional por unidad de tiempo; poseerán, forzosamente, ritmos de acreción elevados, es decir, caerá una cantidad ingente de masa hacia el centro gravitacional. En realidad, el ritmo de acreción de 3C 273 puede deducirse de su potencia emitida, de unos 10^{40} watts. Si el cuásar convierte energía gravitacional en energía radiativa con una eficiencia aproximada del 10 por ciento, el ritmo de acreción será de unos 10^{24} kilogramos por segundo, de unas cuantas masas solares por año. Cada año, vendría a caer en el centro gravitacional de 3C 273 el equivalente de unos cuantos soles.

Para emitir tanta energía, un cuásar ha de tener un núcleo que sea muy compacto y muy masivo. El cuerpo más denso conocido es un agujero negro, objeto cuyas fuerzas gravitacionales son tan grandes que ni la materia ni la radiación pueden escapar del mismo. Según la teoría de la relatividad general, un agujero negro es la configuración estable final de un objeto muy masivo. Por tanto, muchos astrónomos piensan que, sea cual sea el estado original del objeto compacto y masivo de los cuásares, se formará un agujero negro. Por desgracia, nunca se ha hecho una observación directa que permita confirmar o rechazar la existencia de agujeros negros en los cuásares.

¿Por qué parece que algo corre más que la luz?

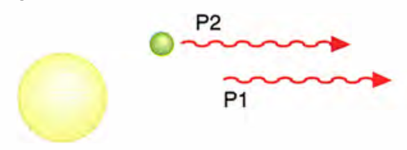
Establece una ley fundamental de la física que la velocidad de la radiación y la materia sólo alcanza la velocidad de la luz, es decir, 300 Mm/s (millones de metros por segundo). Si un trozo de materia es expulsado por un cuásar a velocidades próximas a la de la luz, parecerá, sin embargo, que supere la velocidad de la luz. Este efecto tiene una explicación sencilla. Consideremos un cuerpo que se mueva a 240 Mm/s hacia la Tierra y a 90 Mm/s en una dirección perpendicular a la de observación. (El cuerpo se mueve a 256 Mm/s formando un ángulo de 20 grados con la dirección de observación.)



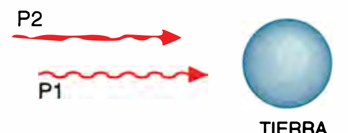
En el momento en que abandona el núcleo, libera un fotón de radiación: fotón 1.



Después de 1 segundo, el fotón 1 ha recorrido 300 Mm y el cuerpo se ha alejado 240 Mm del cuásar hacia la Tierra. Entonces emite un segundo fotón.



Conforme los dos fotones viajan a través del espacio, el fotón 2 permanece 60 Mm detrás del fotón 1 en la dirección paralela a la dirección de observación. Los dos fotones también estarán separados 90 Mm en la dirección perpendicular a la de observación.



El fotón 2 llegará a la Tierra 0,2 segundos después del 1, ya que el retraso temporal ha de igualar la distancia "paralela" (60 Mm) dividida por la velocidad de los fotones (300 Mm/s). La velocidad aparente del cuerpo por el cielo es el cambio "perpendicular" en distancia (90 Mm) dividido por el retraso temporal (0,2 segundos). Por tanto, parece que el cuerpo viaje a 450 Mm/s: el 50 % más deprisa que la velocidad de la luz.

La materia que rodea un agujero negro masivo describirá probablemente una espiral hacia su centro, más que una trayectoria directa; formará, pues, un disco alrededor de la masa central. Este disco de acreción es una estructura que se observa en torno a muchos objetos compactos de nuestra galaxia. Los discos de acreción adquieren tal densidad que la radiación sólo escapa por su superficie, de la misma manera que las estrellas emiten luz sobre todo por la zona superficial llamada fotosfera. Se ha indicado que la luminosidad de la fotosfera de los discos de acreción pudiera ser responsable del gran pico azul observado en la zona visible y ultravioleta del espectro de 3C 273.

La estructura y la dinámica de los cuásares siguen presentando muchos misterios. Los astrónomos desconocen las geometrías de la componente de sincrotrón, de la componente infrarroja y de la de rayos X; tampoco pueden deducir las relaciones que haya entre ellas y los mecanismos de suministro de energía y fenómenos de acreción. Nuestras observaciones han revelado que se trata de problemas complejos; podrían estar muy relacionadas con la generación de los movimientos relativistas. Aunque nuestro conocimiento de los cuásares esté aún lejos de ser completo, esperamos que las observaciones repetidas y las indagaciones teóricas continuarán proporcionando pistas que, como las piezas de un rompecabezas, permitan montar una descripción global.

BIBLIOGRAFIA COMPLEMENTARIA

THEORY OF EXTRAGALACTIC RADIO SOURCES. M. C. Begelman, R. D. Blandford y M. J. Rees en *Reviews of Modern Physics*, vol. 56, n.º 2, Parte 1, págs. 255-351; abril de 1984.

BLACK HOLE MODELS FOR ACTIVE GALACTIC NUCLEI. M. J. Rees en *Annual Review of Astronomy and Astrophysics*, vol. 22, págs. 471-506; 1984.

QUASAR ASTRONOMY. Daniel W. Weedman. Cambridge University Press, 1986.

ACTIVE GALACTIC NUCLEI. T. Courvoisier, R. D. Blandford, L. Woltjer, M. Mayor y H. Netzer. Springer, 1991.

La proteína M de los estreptococos

La bacteria responsable de la faringitis y fiebres reumáticas utiliza esta molécula de la superficie para eludir las defensas corporales. La clave del poder de la proteína reside en su estructura peculiar

Vincent A. Fischetti

Así como los mamíferos se cubren de pelo y las aves de plumas, los microorganismos tienen en su superficie refinadas estructuras para asegurar su supervivencia. Algunas de esas estructuras son moléculas receptoras que dirigen al microbio hacia el nicho específico que constituye su hábitat natural. Otras moléculas protegen al organismo de la destrucción por procesos naturales. Los patógenos que producen enfermedades en las personas son probablemente los microorganismos cuya superficie ha sido mejor estudiada. Virus, parásitos y bacterias están recubiertos de proteínas o moléculas de azúcares que les franquean la entrada en el anfitrión humano, contrarrestando sus defensas.

Una de esas moléculas es la proteína M, sintetizada por ciertas bacterias estreptocócicas. Se pensó en un comienzo que la proteína M tenía una estructura exclusiva. La verdad es que abarca, eso parece, un motivo común, compartido por otras muchas proteínas de superficie bacterianas. Ello convierte a la proteína M en un modelo para otras proteínas de superficie, circunstancia que podría acelerar el desarrollo de algunas terapias antibacterianas. Dado el conocimiento que empezamos a tener de su estructura, se acaricia la idea de prepararla como vacuna contra infecciones estreptocócicas.

VINCENT A. FISCHETTI es profesor y codirector del laboratorio de inmunología y patología bacteriana de la Universidad Rockefeller. Dirigie la revista *Infection and Immunity*. Se formó en el Wagner College y en la Universidad de Long Island, doctorándose por la de Nueva York en 1970. Ha recibido el máximo galardón médico de su país por el trabajo que desarrolla aquí.

Los estreptococos son un grupo de bacterias caracterizadas por su crecimiento concatenado. En forma diversa integran la flora bacteriana habitual en el hombre, sin que revistan especial peligro. *Streptococcus salivarius*, por ejemplo, se encuentra normalmente en la lengua. *Streptococcus mutans* se localiza en la dentadura (y suele estar implicada en la caries dental).

Un subgrupo particular de bacterias estreptocócicas, el grupo A, representado por *Streptococcus pyogenes*, es un patógeno humano. Cada año, y sólo en Occidente, se producen decenas de millones de casos de infecciones estreptocócicas del grupo A; entre ellas, infecciones de la piel y garganta, formas de neumonía y una enfermedad descrita recientemente cuyos síntomas son parecidos a los de un choque tóxico. La infección más común es la faringitis estreptocócica aguda, que afecta principalmente a los niños en edad escolar. Este tipo de faringitis puede considerse un problema sanitario mundial, si se tienen en cuenta las horas de escolaridad y trabajo perdidas, amén de los consiguientes gastos médicos.

El problema se agrava en la garganta. En el 4 por ciento de los casos de faringitis no tratadas, o tratadas mal, la infección estreptocócica puede provocar la fiebre reumática aguda, enfermedad que daña el corazón, concretamente las válvulas cardíacas. La mayoría de las intervenciones quirúrgicas acometidas para corregir válvulas cardíacas dañadas suelen recaer en pacientes que sufrieron fiebres reumáticas durante la infancia. Los daños cardíacos no son causados directamente por las bacterias, sino que son resultado de la respuesta inmune. Los estreptococos engañan al sistema inmune, y hacen que éste ataque a algunos tejidos del propio cuerpo.

Aunque la fiebre reumática no constituya un problema grave en los países avanzados, sí lo es en las naciones en vías de desarrollo. En el Tercer Mundo, el 1 por ciento de los niños de edades comprendidas entre los 5 y los 15 años padecen enfermedades cardíacas reumáticas. En la India, por ejemplo, el problema afecta a casi seis millones de niños en edad escolar.

Los estreptocos del grupo A persisten en los tejidos semanas enteras, gracias fundamentalmente a la proteína M. Este polipéptido confiere resistencia a los estreptocos para evitar su ingestión por parte de los fagocitos humanos, leucocitos que detectan y destruyen a los microorganismos invasores. Rebecca Lancefield, del Instituto Rockefeller (ahora Universidad), descubrió ese fenómeno hace unos 60 años, y se puede demostrar fácilmente con un sencillo experimento. Si se colocan estreptococos en una gota de sangre humana, los fagocitos orillan las bacterias que portan proteína M en su superficie, mostrándose feroces con las que carecen de ella.

¿Qué indica ese comportamiento? Que los estreptococos del grupo A han desarrollado un sistema para evitar algunas de las defensas antimicrobianas humanas. No obstante, es posible resistir a esas infecciones bacterianas si el anfitrión libera anticuerpos contra la proteína M. Esos anticuerpos neutralizan la capacidad protectora de la proteína M y consiguen que los fagocitos engullan y destruyan los estreptococos.

Pero, y eso no es buena noticia, hay más de 80 serotipos, o variedades, diferentes de proteína M. Los ensayos de laboratorio sugieren que los anticuerpos contra un serotipo no protegen contra los demás. En consecuencia, los investigadores piensan que la

exposición a un serotipo estreptocócico de grupo A no inmunizará a un individuo contra posteriores infecciones estreptocócicas de otro tipo. El sistema ideado por los estreptococos para cambiar la molécula M y evitar ser reconocidos por los anticuerpos se denomina variación antigénica. No es exclusivo de los estreptococos. Otros patógenos, incluidos virus y parásitos, utilizan la variación antigénica para evitar que sus moléculas de superficie sean reconocidas por el sistema inmune.

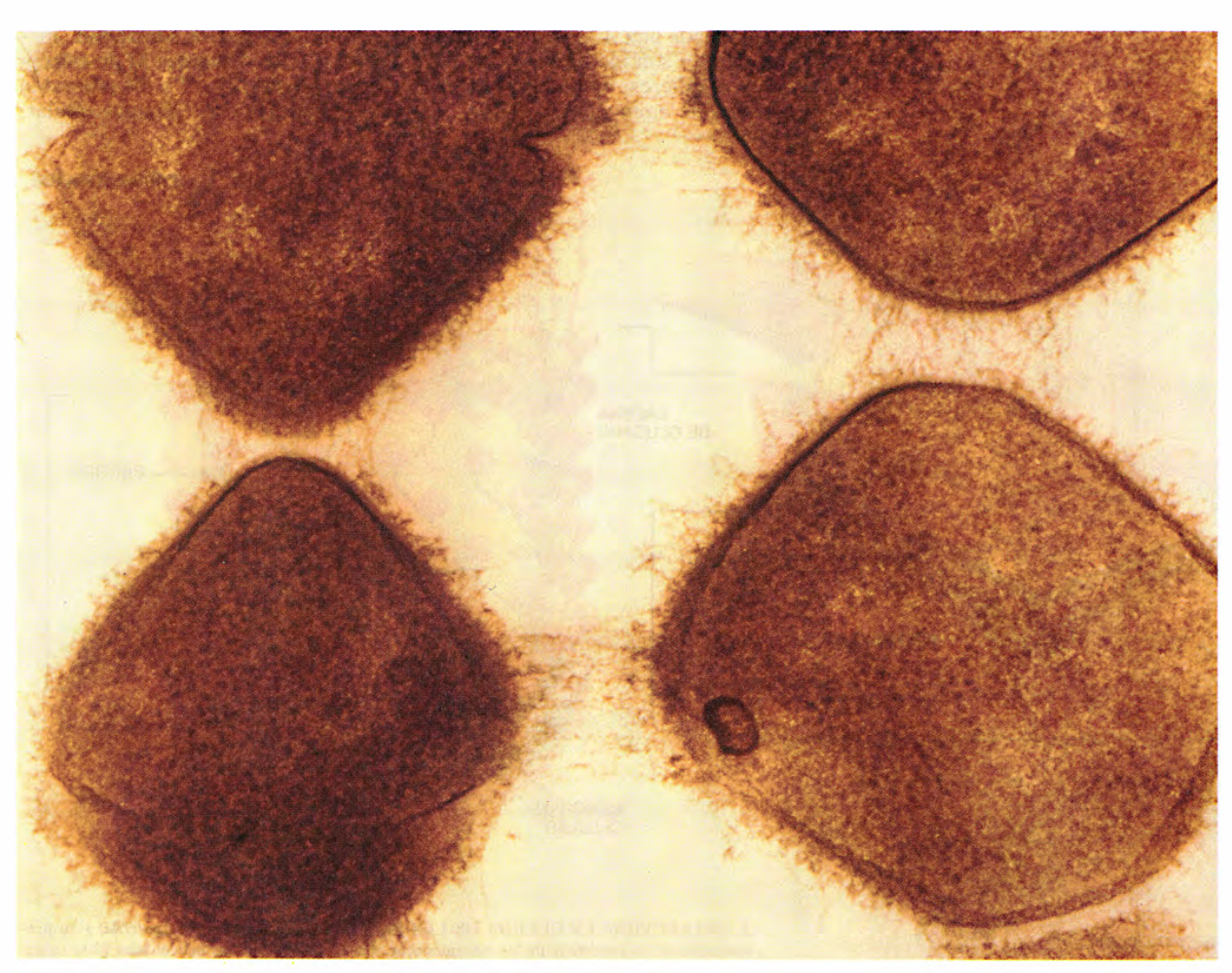
Muchos científicos han trabajado para descifrar el mecanismo implicado en el cambio antigénico de la molécula M y la evasión consiguiente de los estreptococos a la fagocitosis (ataque de los fagocitos). En esa línea viene investigando también mi equipo de la Universidad Rockefeller. La clave parece estar en la estructura de la proteína M, con cuyo concurso la bacteria repele las defensas del cuerpo, al menos de tres maneras. En coherencia con ello, hemos desarrollado ciertas estrategias –y una vacuna potencial– para controlar las infecciones estreptocócicas. De nuestro estudio se desprende cómo abordar la estructura y función de las proteínas de superficie de otros patógenos bacterianos.

Cuando se observa, en el microscopio electrónico, una sección ultrafina de un estreptococo del grupo A, las proteínas M dibujan proyecciones piliformes sobre la superficie de la pared celular bacteriana. La pared celular es una barrera física flexible, aunque fuerte, que protege a la frágil membrana celular. Esta membrana controla el flujo de nutrientes hacia el interior y el de desechos y otros productos hacia fuera. Porta también muchas enzimas necesarias para el metabolismo celular y la síntesis de proteínas. En la mayoría de las bacterias, ligada a la superficie externa de la pared celular existe una capa de carbohidratos complejos.

Los estreptococos son bacterias gram-positivas, caracterizadas por tener una pared celular gruesa, que en un determinado procedimiento de identificación retiene un colorante químico específico. (La técnica de tinción fue desarrollada por el bacteriólogo danés H.C.J. Gram.)

La pared celular de los estreptococos está formada por largas cadenas lineales; estas unidades de glucano se componen de N-acetilglucosamina-N-acetil-murámico, un azúcar disacárido. Las cadenas se conectan entre sí por péptidos breves (cadenas de aminoácidos); de ahí que se llame peptidoglucano a la pared celular. El peptidoglucano de un organismo gram-positivo consta de unas diez capas de cadenas de glucano interconectadas por péptidos, formando todo ello una malla espesa. El componente peptídico varía a veces entre diferentes especies gram-positivas, pero la estructura fundamental de los peptidoglucanos es casi siempre similar.

Un primer paso importante en el desciframiento de la estructura de la



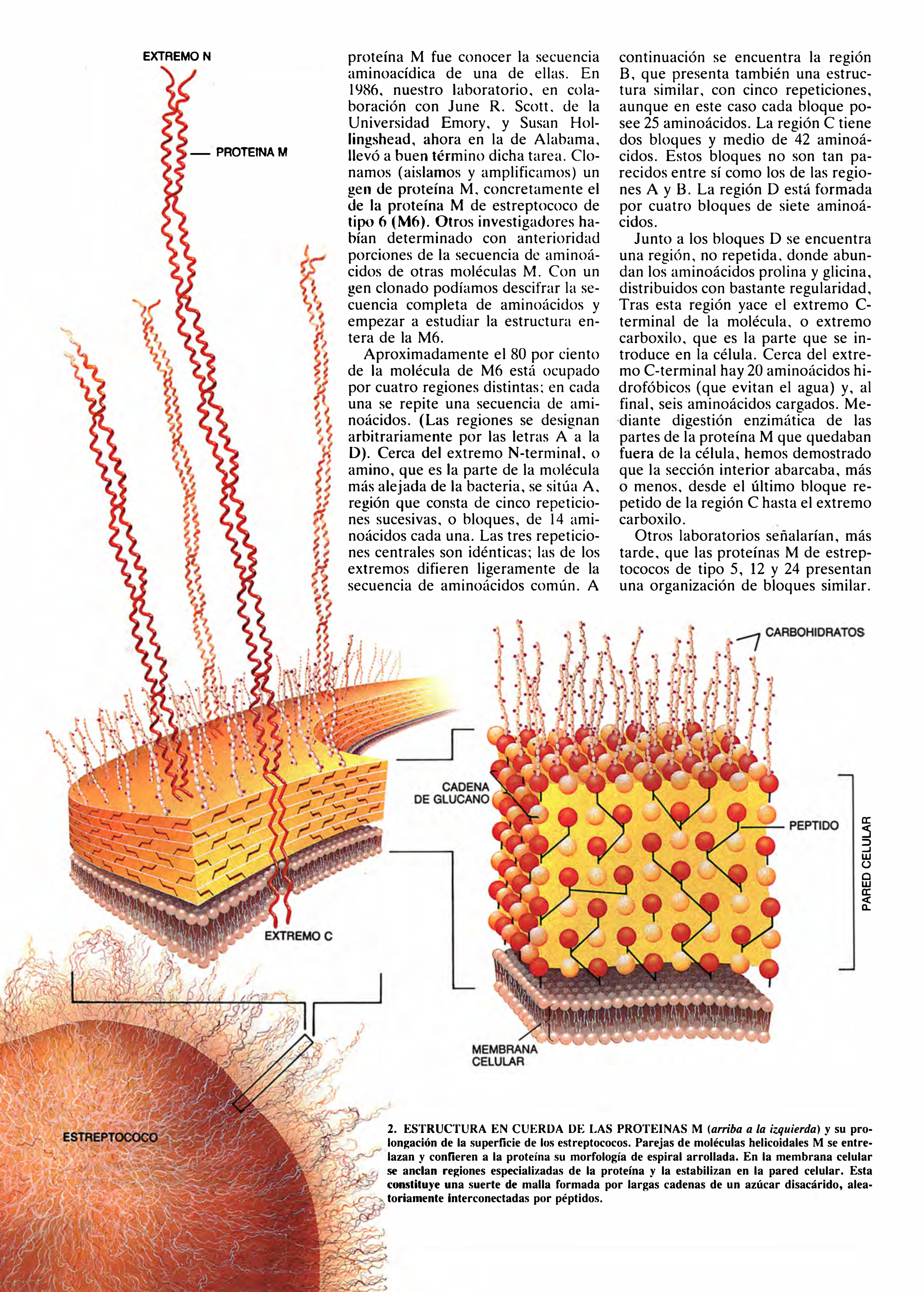
1. DE PROTEINA M son los largos filamentos piliformes que sobresalen de la superficie de los estreptococos del grupo A. Ayudan a las bacterias a evitar las defensas inmunológicas de sus hospedadores, los humanos, e impedir así su eliminación del cuerpo. Estudios recientes sobre la proteína M estreptocócica revelan las características estructurales específicas de la molécula y el modo de cumplir su misión.

proteína M fue conocer la secuencia aminoacídica de una de ellas. En 1986, nuestro laboratorio, en colaboración con June R. Scott, de la Universidad Emory, y Susan Hollingshead, ahora en la de Alabama, llevó a buen término dicha tarea. Clonamos (aislamos y amplificamos) un gen de proteína M, concretamente el de la proteína M de estreptococo de tipo 6 (M6). Otros investigadores habían determinado con anterioridad porciones de la secuencia de aminoácidos de otras moléculas M. Con un gen clonado podíamos descifrar la secuencia completa de aminoácidos y empezar a estudiar la estructura entera de la M6.

Aproximadamente el 80 por ciento de la molécula de M6 está ocupado por cuatro regiones distintas: en cada una se repite una secuencia de aminoácidos. (Las regiones se designan arbitrariamente por las letras A a la D). Cerca del extremo N-terminal, o amino, que es la parte de la molécula más alejada de la bacteria, se sitúa A, región que consta de cinco repeticiones sucesivas, o bloques, de 14 aminoácidos cada una. Las tres repeticiones centrales son idénticas; las de los extremos difieren ligeramente de la secuencia de aminoácidos común. A continuación se encuentra la región B, que presenta también una estructura similar, con cinco repeticiones, aunque en este caso cada bloque posee 25 aminoácidos. La región C tiene dos bloques y medio de 42 aminoácidos. Estos bloques no son tan parecidos entre sí como los de las regiones A y B. La región D está formada por cuatro bloques de siete aminoácidos.

Junto a los bloques D se encuentra una región, no repetida, donde abundan los aminoácidos prolina y glicina, distribuidos con bastante regularidad. Tras esta región yace el extremo C-terminal de la molécula, o extremo carboxilo, que es la parte que se introduce en la célula. Cerca del extremo C-terminal hay 20 aminoácidos hidrofóbicos (que evitan el agua) y, al final, seis aminoácidos cargados. Mediante digestión enzimática de las partes de la proteína M que quedaban fuera de la célula, hemos demostrado que la sección interior abarcaba, más o menos, desde el último bloque repetido de la región C hasta el extremo carboxilo.

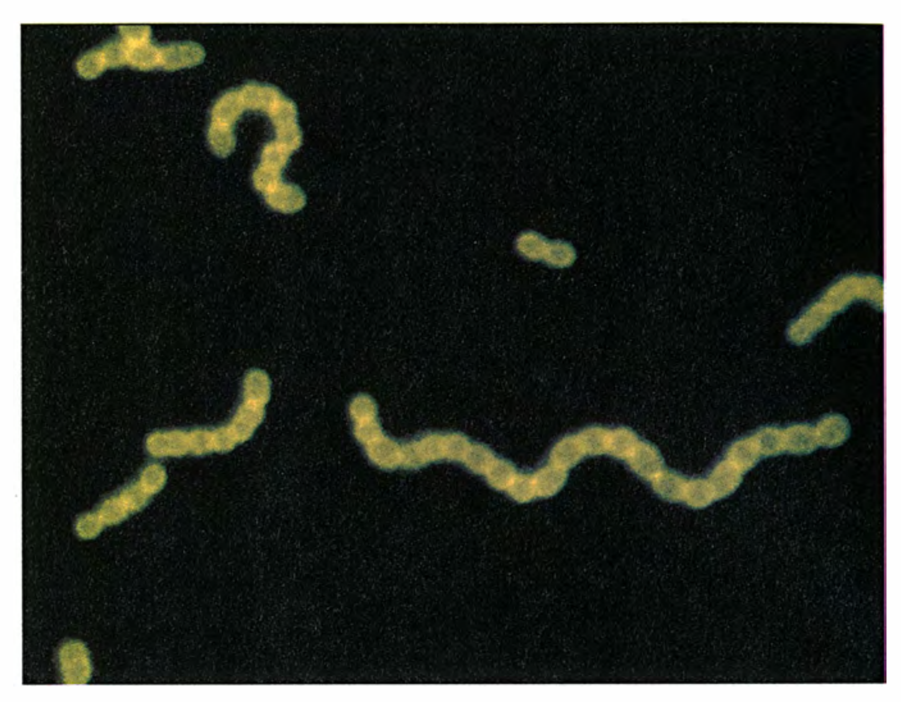
Otros laboratorios señalarían, más tarde, que las proteínas M de estreptococos de tipo 5, 12 y 24 presentan una organización de bloques similar.



2. ESTRUCTURA EN CUERDA DE LAS PROTEINAS M (*arriba a la izquierda*) **y su prolongación de la superficie de los estreptococos. Parejas de moléculas helicoidales M se entrelazan y confieren a la proteína su morfología de espiral arrollada. En la membrana celular se anclan regiones especializadas de la proteína y la estabilizan en la pared celular. Esta constituye una suerte de malla formada por largas cadenas de un azúcar disacárido, aleatoriamente interconectadas por péptidos.**

Si se alinean las secuencias aminoacídicas de esas proteínas M, se comprueba que los extremos carboxilos comparten una homología de más del 98 por ciento. En las cercanías del extremo N, sin embargo, aumentan las diferencias entre las distintas proteínas M. Los bloques A y una breve región de 11 aminoácidos del extremo N son específicos de cada molécula M. Sabemos ahora que muchas proteínas de superficie de bacterias gram-positivas contienen bloques de secuencias repetidas.

Cuando mi colega Belur M. Manjula y yo analizamos con más detenimiento la secuencia de aminoácidos de la M6, descubrimos otro detalle estructural interesante. A lo largo de todos lo bloques de repeticiones persistía un motivo insólito de siete aminoácidos: hidrofóbicos los de la primera y cuarta posición y los demás permitían el giro de la proteína sobre sí misma, a modo de espiral, adoptando la configuración en hélice alfa. Esta observación nos movió a buscar en la bibliografía otras proteínas con características similares. Todas las moléculas que encontramos tenían hélices alfa y una conformación en espiral plegada, esto es, dos cadenas de hélices alfa arrolladas sobre sí mismas, a modo de cuerda.



3. LAS BACTERIAS ESTREPTOCOCICAS (*Streptococcus pyogenes*) crecen formando cadenas; provocan fiebres reumáticas. En esta micrografía se han marcado los estreptococos con un colorante fluorescente, con anticuerpos monoclonales que se unen a la proteína M.

La estructura en bobina, o espiral arrollada, se forma gracias a la geometría de la hélice alfa y a las propiedades de los aminoácidos constituyentes. Para dar una vuelta de hélice alfa se precisan 3,6 aminoácidos. Por tanto, en la proteína M6 los aminoácidos hidrofóbicos de las posiciones 1 y 4 quedan alineados, formando una banda inclinada a lo largo de la hélice. Por su naturaleza hidrofóbica, esos aminoácidos prefieren quedar ocultos en la proteína, sin entrar en contacto con las moléculas de agua del ambiente. Las hélices alfa de M6 se arrollan dos a dos, para encerrar o proteger sus aminóacidos hidrofóbicos.

Este motivo de siete unidades que se repetía en la secuencia de aminoácidos de M6 confería a la molécula la apariencia de una larga varilla helicoidal. El patrón de la M6 ni es perfecto, ni se encuentra en otras muchas estructuras abobinadas. Tales irregularidades explican probablemente la flexibilidad de las moléculas M, observada en las electromicrografías. Más aún. Las características de tales irregularidades difieren en las regiones repetidas A, B y C. De lo que cabe inferir que cada región repetida evolucionó de manera independiente y podría tener una función distinta. Así parecen confirmarlo algunos experimentos en marcha para comprobar la validez de la hipótesis.

En colaboración con Carolyn Cohen, de la Universidad Brandeis, y George N. Phillips Jr., ahora en la Universidad Rice, acumulamos más datos físicos y químicos sobre la M6. Nuestras medidas confirmaban que cada fibra de proteína M6 insertada en la pared celular del estreptococo medía unos 50-60 nanómetros (1 metro = mil millones de nanómetros) y, también, que constaba de una sola molécula en bobina.

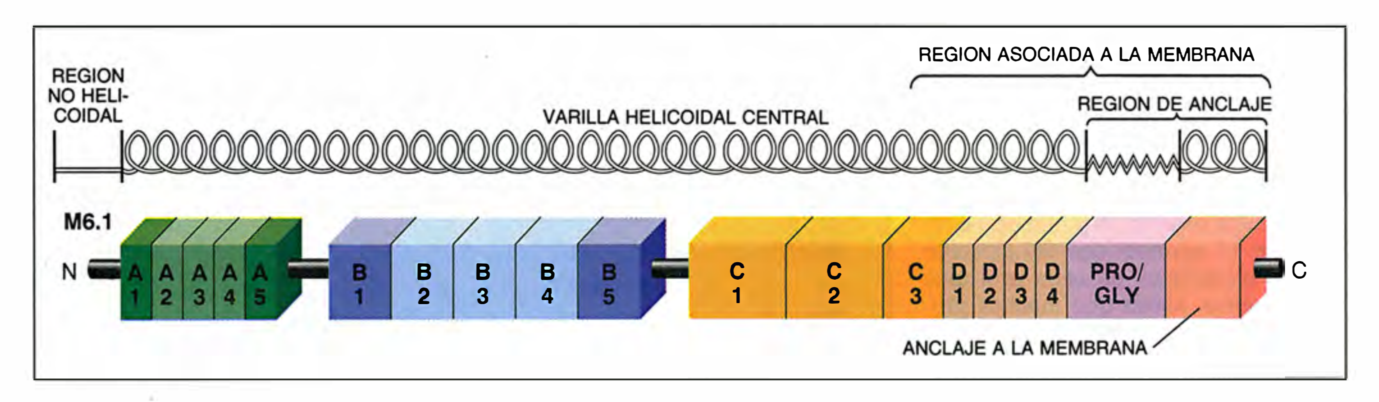
Es probable que las proteínas M de todos los serotipos estén construidas siguiendo un esquema básico similar. Una larga región, a modo de huso de bobina, flanqueda esta última por una sección colgante en el extremo N y otra de anclaje en el extremo C. Porque la estructura abobinada de hélices alfa puede acomodar un gran número de secuencias de aminoácidos distintas, se pueden construir muchas proteínas M diferentes con la misma conformación general. Las presiones de la selección natural sobre la molécula parecen haber favorecido esa figura de carrete o bobina. Los estreptococos mutantes que fabricasen proteínas M con una estructura alejada de ese patrón general, no sobrevivirían probablemente en el cuerpo humano.

Para que la proteína M proteja al estreptococo, debe anclarse en dicho organismo. No se conoce bien el mecanismo que afianza las proteínas de superficie en las bacterias gram-positivas; algo, sin embargo, hemos avanzado con nuestro estudio sobre la proteína M.

Con mi colega de la Rockefeller Vijaykumar Pancholi, comprobé que se podía eliminar la pared celular de un estreptococo utilizando una enzima que corta al peptidoglucano. Si hacíamos eso en una solución de azúcar rafinosa al 30 por ciento, la alta presión osmótica de la solución impedía que se rompiese la membrana de la bacteria. En esas condiciones, la proteína M seguía unida a la membrana celular. Este sencillo experimento nos demostraba que la proteína M estaba unida a la membrana celular, y no a la pared celular, como muchos investigadores habían sospechado.

Pensábamos que los 20 aminoácidos hidrofóbicos próximos al extremo C terminal estarían insertados en la membrana hidrofóbica, mientras que los aminoácidos cargados situados más en el extremo se introducirían en el citoplasma acuoso. Dado que los aminoácidos cargados podrían resistir moviéndose en un ambiente hidrofóbico, serían como nudos al final de una cuerda, es decir, impedirían que la molécula M se soltase de la membrana.

Este mecanismo podría ser válido para algunas proteínas que se anclan en la membrana. Pruebas más recientes indican, sin embargo, que el mecanismo de inserción utilizado por la proteína M y otras proteínas bacterianas de superficie podrían encerrar mayor complejidad. Nuestros estu-



4. SECUENCIA PROTEINICA de la molecula M6, determinada tras clonar su gen. Aproximadamente el 80 por ciento de la proteína M está formado por bloques de secuencias de aminoácidos repetidas (A, B, C y D). La región rica en prolinas y glicinas ayuda a que la proteína atraviese la pared celular. Los aminoácidos próximos al extremo carboxilo permiten su inserción en la membrana celular.

dios revelan que todas las proteínas de superficie de las bacterias grampositivas poseen una organización similar de aminoácidos hidrofóbicos y cargados en su extremo carboxilo. Más. Una breve secuencia de seis aminoácidos, próxima a la región hidrofóbica, persiste en todas las proteínas de superficie conocidas de las bacterias gram-positivas. La secuencia se enhebra con leucina, prolina, serina, treonina, glicina y ácido glutámico; se la designa con la abreviatura LPSTGE.

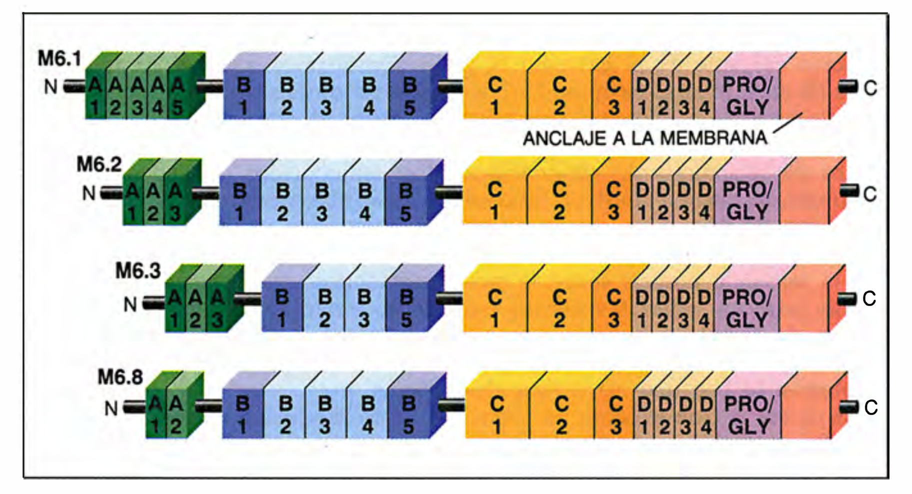
La importancia de la secuencia LPSTGE en la inserción de la proteína M (y probablemente en todas las demás proteínas con esa misma secuencia) se puso de manifiesto en los experimentos genéticos realizados por Olaf Schneewind en mi laboratorio. Observó que, cuando eliminaba la secuencia LPSTGE del gen de la proteína M, la molécula producida era incapaz de insertarse en la membrana bacteriana. ¿Qué nos decía ese resultado? Dos cosas sugería: que, para el anclaje en la membrana, no bastaban el dominio hidrofóbico y los aminoácidos cargados del extremo C, y que el motivo LPSTGE podía ser una señal decisiva de arranque del proceso. La naturaleza de la señal y el mecanismo de inserción son objeto de intensas investigaciones.

En casi todas las proteínas de superficie encontradas en bacterias gram-positivas, existe otra región característica que abarca unos 50 o 75 aminoácidos en el lado N-terminal de la región hidrofóbica. Esta parte se ubica probablemente en el peptidoglucano, y predominan en la misma los aminoácidos prolina, glicina, treonina y serina. La razón de ese predominio no está muy clara, pero es sabido que prolinas y glicinas pueden crear vueltas y ángulos en las proteínas. Según una hipótesis, se puede tejer un entramado de interconexiones en el peptidoglucano, a través de las curvaturas inducidas por prolinas y glicinas, estabilizando así la posición de la proteína M en la pared celular.

El descubrimiento de que todas las proteínas de superficie de bacterias gram-positivas se insertan utilizando un mecanismo similar abre nuevas perspectivas en el control de las infecciones causadas por esos organismos. Las proteínas de superficie facilitan la labor infecciosa de los organismos patógenos. Si se impide que las proteínas se anclen en la bacteria, se evitarán las infecciones y se orillarán problemas asociados con la resistencia a las terapias basadas en antibióticos.

Lo mismo que las estructuras del extremo C terminal de la molécula nos ilustran sobre la inserción de la proteína M en la bacteria, así también las estructuras del extremo N terminal nos permiten inducir el papel de las moléculas en el rechazo a los fagocitos. El extremo N terminal de todas las moléculas M tiene un exceso de aminoácidos dotados de carga negativa; hay, en efecto, una carga neta negativa en esa región. Las células de mamíferos presentan también una carga neta negativa en su superficie. La carga de la proteína M puede, por tanto, haber evolucionado para impedir el contacto entre estreptococos y células fagocitarias, gracias a la repulsión electrostática. Parece verosímil que una de las funciones de la varilla central de la proteína M sea la de actuar como una suerte de tallo que mantiene alejado, de la superficie de la bacteria, el extremo N terminal cargado negativamente y a los fagocitos.

En el extremo N terminal de la varilla arrollada existe una región hipervariable. Esta parte de la molécula tiene una secuencia característica en cada serotipo M. La región hipervariable está formada por una secuencia no helicoidal de 11 aminoácidos y el bloque de repeticiones A. Hemos descubierto que la región hipervaria-



5. DIFERENTES FORMAS DE PROTEINAS M, surgidas en los estreptococos mutantes por deleción de algunas de las secuencias repetidas que se encuentran en las moléculas parentales. La variación antigénica que resulta de estos cambios facilita que las bacterias se hurten a la acción del sistema inmune. La mayoría de las variaciones estructurales de la proteína M se producen cerca del extremo amino. La mitad C terminal de la molécula es menos variable.

ble desempeña un papel importante en la actividad biológica de la molécula. Sólo los anticuerpos contra esa región dejan al estreptococo indefenso ante los fagocitos. Esta observación explica por qué sólo los anticuerpos específicos de serotipo protegen contra las infecciones estreptocócicas.

Así se enrolla la proteína M

La secuencia de la proteína M tiene un patrón regular de aminoácidos hidrofóbicos.

Cuando la secuencia adopta la conformación en hélice alfa, se forma una banda hidrofóbica.

Parejas de cadenas helicoidales se arrollan para proteger sus regiones hidrofóbicas.

Durante años, los investigadores se han preguntado por el proceso en cuya virtud los estreptococos cambian las secuencias de sus regiones hipervariables. Nuestro descubrimiento del carácter variable, a tenor de los organismos, del tamaño de las moléculas M ofrece una explicación parcial. Examinamos de entrada 20 serotipos de proteínas M recogidos de infecciones estreptocócicas durante un período de 40 años. Las moléculas presentaban una amplia gama de tamaños, de 41.000 a 80.000 unidades de peso molecular. Detectamos una variación similar entre las proteínas M6 estreptocócicas tomadas de varios niños de un centro asistencial durante una epidemia de faringitis.

Para determinar con qué rapidez cambiaba de tamaño la molécula, examinamos la proteína M de muestras de una estirpe de estreptococos M6, desarrollada en nuestro laboratorio durante más de 30 años. Por sorprendente que parezca, no había pruebas de que la proteína M hubiese cambiado de tamaño durante ese tiempo.

Analizamos también otras estirpes estreptocócicas recogidas periódicamente a pacientes durante varias semanas, en los años cuarenta, antes de la generalización del empleo de la penicilina. Los estreptococos aislados del mismo paciente con un intervalo semanal presentaban proteínas M del mismo serotipo, aunque de diferente tamaño. En la garganta del paciente, así pensamos, la presión de la selección natural sobre los estreptococos favorecía la aparición de mutantes que producían proteínas M de diferente tamaño. Presiones que no sufría la cepa que cultivábamos en el laboratorio.

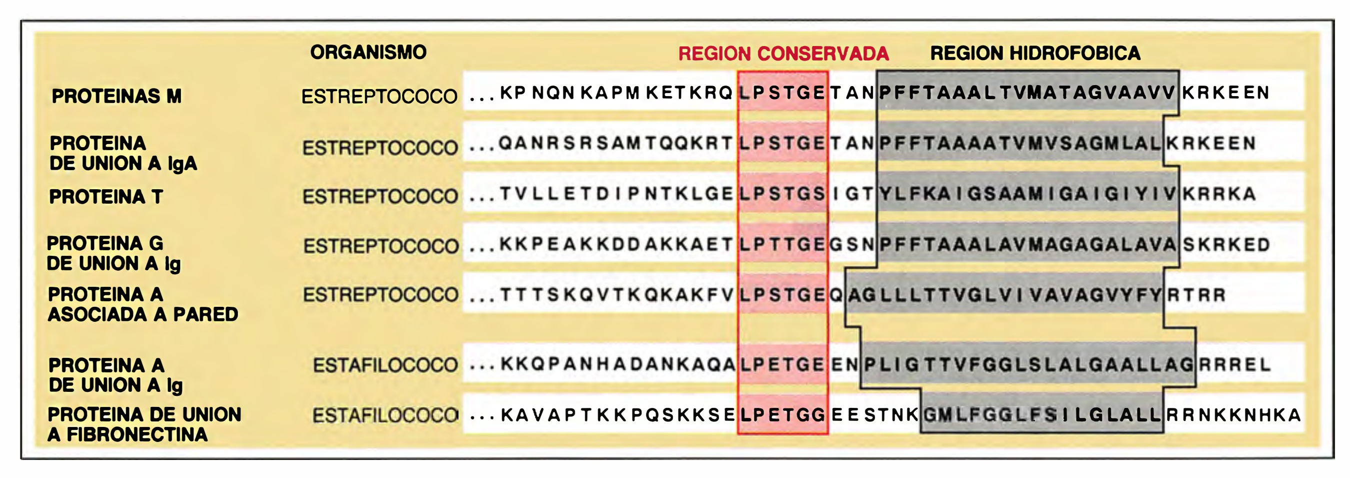
La investigación posterior nos reveló la presencia, en número limitado, de mutaciones que afectaban al tamaño entre las células de la estirpe M6 de laboratorio, de apariencia homogénea. Según nuestros cálculos, los mutantes que fabrican proteínas M más pequeñas surgen una vez cada 2000 cadenas estreptocócicas. Esta cadencia es demasiado alta para ser explicada por mutaciones puntuales espontáneas, esto es, cambios en una de las bases del ADN. Cambios de este tipo se esperan con una frecuencia de uno cada millón o 10 millones de cadenas.

Para examinar la causa de este cambio de tamaño, analizamos los genes determinantes de la proteína M de mutantes distintos. Y así demostramos que los cambios se debían a la pérdida de secuencias de ADN de ciertos bloques de repeticiones de la proteína M: un mutante tenía una deleción que afectaba a dos bloques de tipo A, y otro había eliminado uno de los bloques B; en un tercer caso, la deleción abarcaba desde el centro del primer bloque A hasta el centro del tercer bloque A. Como los finales de los bloques son ligeremente distintos, la deleción provocaba un cambio en la secuencia de aminoácidos. Salía a la luz de ese modo el origen de los cambios de las proteínas M. Durante la replicación del ADN en los estreptococos que se multiplicaban, se producían errores ocasionales que eliminaban segmentos de ADN de algunos bloques. Como consecuencia de ello, la descendencia mutante producía moléculas M menores. En la naturaleza aparecen moléculas M6 más largas, pero con nuestras técnicas de laboratorio no podemos aislarlas.

Apoyados en ese hallazgo, mi colega Kevin Jones y yo nos preguntamos a continuación si los anticuerpos que reaccionaban contra una proteína M parental lo harían también con las moléculas M mutantes, más cortas. Para comprobarlo, utilizamos anticuerpos específicos contra las regiones repetidas A y B de una proteína M6 parental. Y vimos que los anticuerpos no se unían a las proteínas mutantes M o lo hacían sólo muy débilmente, si sus sitios normales de unión estaban en una de las áreas delecionadas o en otra inmediata.

Mediante la modificación del tamaño, las proteínas M evitan la unión de los anticuerpos a la molécula. Los estreptococos aprovechan los errores genéticos para cambiar el tamaño y el carácter antigénico de sus moléculas M, escapando así del reconocimiento y destrucción por parte del sistema inmune del hospedador. Los estreptococos mutantes que aparecían en los cultivos semanales de los pacientes con faringitis, habrían sobrevivido porque sus proteínas M no eran las dianas de los anticuerpos fabricados contra los organismos parentales.

Estos hallazgos ponen sobre la mesa la estrategia utilizada por los organismos patógenos para sobrevivir en el ambiente. Le sacan partido a su capacidad de replicación rápida. Un organismo que se divida cada 30 minutos producirá más de un millón de células hijas en 10 horas. Los errores producidos durante este proceso replicativo crean organismos mutantes. Normalmente, un mutante no sobrevivirá, ya que su defecto se produce en una función vital para el crecimiento. En otros casos, sin embargo, el cambio genético puede permitir que el mutante sobreviva y prospere



6. REGIONES SIMILARES presentes cerca del extremo C de varias proteínas de diversos grupos de bacterias gram-positivas. Cada letra representa un aminoácido de la secuencia proteínica. Las homologías en el extremo C sugieren la existencia de un mecanismo común para la inserción de los extremos C de las proteínas M y de esas restantes moléculas en las bacterias. (Ilustración de Edward Bell.)

en un ambiente no apto para otras células hijas.

El mismo principio se aplica a la proteína M y otras moléculas de superficie con capacidad de transformación. Los cambios suelen producirse en zonas genéticas "calientes", donde predominan secuencias repetidas. Por tratarse de regiones con particular tendencia al error durante la replicación del ADN, los microorganismos mantienen las secuencias repetidas para asegurar un suministro estable de mutantes. Aunque este mecanismo explica cómo ocurren los cambios en las regiones repetidas, no aporta información alguna sobre el proceso a través del cual cambian los 11 aminoácidos del extremo amino entre los distintos serotipos.

De dos trucos podría valerse, hemos dicho, la proteína M para confundir a las defensas de un hospedador humano: la repulsión N terminal y la variación antigénica. Para entender un tercer mecanismo necesitamos revisar primero otro sistema de supervivencia, mediante el cual el hospedador identifica y neutraliza la infección bacteriana.

Los mamíferos portan en su sangre y fluidos linfáticos una disposición compleja de proteínas y enzimas –el sistema del complemento– que reconoce a los microorganismos patógenos y a otras partículas foráneas. Cuando una partícula foránea entra en el cuerpo, la proteína del complemento C3b rápidamente se une a ella. La C3b marca al patógeno como algo que debe ser aniquilado y expulsado del cuerpo por los leucocitos u otros sistemas parecidos. Un proceso tan rápido para reconocer invasores puede cometer errores y marcar equivocadamente células normales de hospedador (mamífero). Para evitarlo, el suero sanguíneo contiene una serie de proteínas reguladoras que pueden deshacer, o bloquear, la unión de la C3b a esas células.

Mientras se investigaba el motivo de que los estreptococos no quedaran eliminados tras una infección, se observó que la C3b se unía eficazmente a los estreptococos carentes de proteína M; lo que no ocurría con los que sí la llevaban. En colaboración con Rolf Horstman, del Instituto Bernhard-Nocht, de Hamburgo, descubrimos la razón de ese distinto comportamiento. El factor H, una de las proteínas reguladoras del sistema del complemento, se trababa específicamente con la proteína M. Engarce que limita la interacción de C3b con el estreptococo y la protege de la eliminación.

En cierto sentido, los estreptococos que portan moléculas M inteligentemente se disfrazan de células humanas, para evadir el sistema del complemento. Más tarde, comprobamos que el factor H se unía a la proteína M por la región conservada del bloque de tipo C.

Esa estrategia para evitar ser barrido del cuerpo ha hecho del estreptococo un parásito humano con éxito. De ello inferimos que, en el curso de una infección, el factor H se une a la proteína M, impidiendo el marcado destructor del estreptococo. Si la infección no se ataja, el hospedador generará anticuerpos contra la proteína M en el intervalo de una a dos semanas.

Creemos que los anticuerpos que se unen a los bloques B y C de la molécula caen bajo la influencia del factor H. En consecuencia, esos anticuerpos no contribuyen a eliminar la infección. Sólo los anticuerpos contra la región hipervariable cercana al extremo N terminal están lo suficientemente lejos del factor H como para poner en marcha la fagocitosis de las bacterias mediada por anticuerpos. Los anticuerpos contra el extremo N neutralizan también la carga negativa de la región y ayudan, por tanto, a la fagocitosis.

El ciclo completo de infección, producción de anticuerpos y limpieza del cuerpo dura unos 14 días. En ese intervalo, el estreptococo puede generar un gran número de descendientes, algunos de los cuales serán idénticos a sus padres, pero otros serán mutantes que han cambiado la proteína M. Las células que tienen éxito suelen transmitirse a otros individuos, donde comienzan un nuevo ciclo, que perpetúa la especie.

Nuestros estudios estructurales han ayudado a entender de qué modo una infección estreptocócica de grupo A puede causar indirectamente el daño cardíaco asociado con la fiebre reumática. Comparamos las secuencias de aminoácidos de la proteína M con las de otras proteínas almacenadas en una base de datos informatizada. La proteína M tenía una homología del 48 por ciento con otras proteínas fibrosas, portadoras asimismo de espirales arrolladas: tropomiosina, miosina y queratina, todas ellas presentes en tejidos de mamíferos. Esta similitud acarrea consecuencias inmunológicas.

Característica de la fiebre reumática es la presencia, en el suero sanguíneo del paciente, de anticuerpos que reaccionan contra el tejido muscular, en particular el tejido cardíaco. No es normal que se fabriquen anticuerpos contra los tejidos propios. Los investigadores han descubierto, sin embargo, que los anticuerpos con reacción cruzada pueden inducirse a veces por una molécula del organismo infeccioso que remede otra del ma-

mífero hospedador. Durante la fabricación de anticuerpos contra las moléculas microbianas, para limpiar la infección, se le engaña al cuerpo para que sintetice anticuerpos contra sus propios tejidos, un proceso lesivo en potencia.

Edwin H. Beachney y sus colaboradores, de la Universidad de Tennessee en Menfis, observaron que los conejos inmunizados con proteínas M purificadas producían anticuerpos contra las proteínas M y contra la miosina, una proteína muscular de mamíferos. Algunos de esos anticuerpos reaccionaban también contra la tropomiosina y otras proteínas musculares. Porque esas proteínas son también moléculas espirales y arrolladas, es probable que los anticuerpos con reacción cruzada reconozcan una característica conformacional común de la proteína M y de las moléculas de mamíferos. Se desconoce, sin embargo, el papel exacto de tales anticuerpos de reacción cruzada en la génesis de la enfermedad cardíaca reumática.

Se ha intentado aplicar los conocimientos sobre la estructura y función de la proteína M al desarrollo de una vacuna que impida las infecciones estreptocócicas. Los anticuerpos contra la región hipervariable N terminal iniciarían la fagocitosis; ahora bien, y lo dijimos antes, las vacunas preparadas a partir de esa región sólo protegen contra un tipo de estreptococo. Tales vacunas deberían incluir, por tanto, muchas secuencias N terminales diferentes. Ese problema, junto a la naturaleza en cambio constante de la región N terminal, resta atractivo a semejante aproximación.

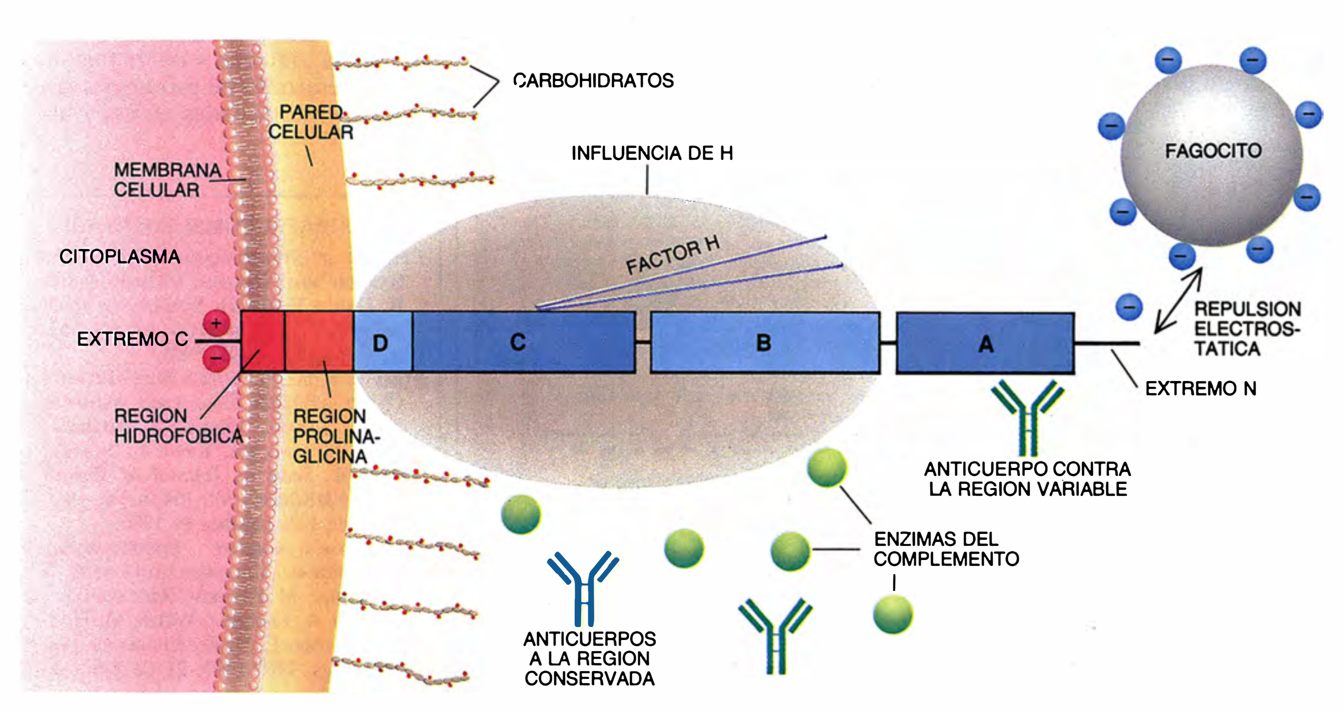
Hemos explorado una estrategia alternativa. Amén de apoyarse en las regiones conservadas de la molécula M que no están ocultas en la pared celular o en la membrana, la táctica se inspira en la incidencia de la faringitis estreptocócica de tipo A, que suele mostrar su máximo entre los 6 y 7 años, para disminuir rápidamente después de los 10. La incidencia más baja de faringitis se da a partir de los 20 años. Es improbable, por tanto, que los adultos se hayan expuesto al centenar casi de tipos de estreptococos, como se ha corroborado en experimentos recientemente realizados en mi laboratorio.

La menor incidencia de faringitis estreptocócica en los adultos puede tener su origen en algún factor no definido, y relacionado con la edad, o tal vez en la existencia de un grado más amplio de inmunidad, adquirida por los individuos a través de los contactos con estreptococos durante la infancia. Creemos que la protección contra la faringitis podría estar inducida por anticuerpos contra algunas regiones conservadas en diferentes serotipos. De esa manera, la exposición a los estreptococos durante la infancia podría permitir el desarrollo de un repertorio de anticuerpos contra las regiones conservadas de la proteína M; son esos anticuerpos los que confieren protección en una fase posterior de la vida.

Para comprobar esta hipótesis, mi colega Debra Bessen y yo preparamos péptidos sintéticos de una veintena de aminoácidos. Cada péptido era copia de alguna secuencia de aminoácidos encontrada en la región C conservada de la proteína M. Los péptidos se unieron con una subunidad inocua de la toxina del cólera. Subunidad que sirvió para transportar los péptidos pequeños y ayudarles a estimular una respuesta inmune contra ellos. Se aplicó el complejo, en aerosol, por vía nasal a los ratones, varias veces durante un período de dos semanas; y de nuevo al cabo de tres semanas, a modo de refuerzo. Esperamos un tiempo para que los ratones desarrollaran anticuerpos apropiados contra el péptido, y los sometimos a exposiciones intranasales y orales de estreptococos vivos del grupo A.

Los animales que recibieron el péptido-vacuna fueron significativamente menos propensos a la faringitis estreptocócica que sus compañeros sin tratar. En esencia, las vacunas prevenían a los ratones contra las afecciones de garganta. Los anticuerpos contra las regiones conservadas de la proteína M bloqueaban, eso entendemos, el ataque y colonización de la faringe (tejido de la garganta) por parte de los estreptococos. Observamos el efecto protector aun cuando el organismo patógeno poseía un serotipo M diferente del utilizado en la vacuna.

Para ensayar un enfoque distinto, en colaboración con Dennis E. Hruby, de la Universidad estatal de



7. ENGAÑAR AL SISTEMA INMUNE es el principal objetivo de la proteína M. Las cargas negativas del extremo N pueden repeler a los fagocitos. La unión del factor H, una proteína reguladora producida por el hospedador humano, protege las regiones más conservadas de la proteína M de la acción de los anticuerpos y enzimas del complemento. Sólo los anticuerpos contra la región hipervariable pueden eliminar una infección.

Oregón, transferimos el gen de la región conservada de la proteína M al virus de la viruela de vacas. (Estos virus, base de la vacuna contra la viruela, son débiles y bastante inocuos para el hombre.) Cuando se infectaban células de mamíferos con los virus genéticamente alterados, las células producían la región conservada de la proteína M. Vacunamos a los ratones intranasalmente con esos virus; quedaban así inmunizados contra los estreptococos vivos.

En resumen, parece posible sensibilizar al sistema inmune de los animales contra la región conservada de la proteína M. Impidiendo la etapa de colonización inicial de una infección, nos ahorraremos el empleo de anticuerpos específicos de tipo, que serían necesarios en el caso exclusivo de que los estreptococos colonizaran e invadieran los tejidos.

Desde que Lancefield identificó la proteína M como un determinante importante de la virulencia de los estreptococos, los científicos han tardado casi 60 años en desarrollar formas de protección contra las infecciones. Pasarán varios años más, probablemente, antes de que se pueda disponer de una vacuna para el hombre. El progreso ha sido posible gracias al conocimiento de la estructura molecular de la proteína M, la localización de sus dominios funcionales y el método por el cual la proteína repele el ataque fagocitario. Es apasionante comprobar que los estudios sobre la proteína M no sólo beneficiarán a las personas con riesgo de infecciones estreptocócicas y fiebres reumáticas, sino que servirán también para el desarrollo de estrategias contra otras enfermedades víricas y bacterianas.

BIBLIOGRAFIA COMPLEMENTARIA

EPITOPES OF STREPTOCOCCAL M PROTEINS SHARED WITH CARDIAC MYOSIN. James B. Dale y Edwin H. Beachey en *Journal of Experimental Medicine*, vol. 162, n.º 2, págs. 583-591; 1985.

STREPTOCOCCAL M PROTEIN SIZE MUTANTS OCCUR AT HIGH FREQUENCY WITHIN A SINGLE STRAIN. Vincent A. Fischetti, Mary Jarymowycz, Kevin F. Jones y June R. Scott en *Journal of Experimental Medicine*, vol. 164, n.º 4, págs. 971-980; 1 de octubre de 1986.

PROTECTION AGAINST STREPTOCOCCAL PHARYNGEAL COLONIZATION WITH A VACCINIA: M PROTEIN RECOMBINANT. Vincent A. Fischetti, Walter M. Hodges y Dennis E. Hruby en *Science*, vol. 244, págs. 1487-1490; 23 de junio de 1989.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Ciencia y sociedad

¿Vacuna contra la malaria?

¿Habrá que agradecer a los ratones el descubrimiento de una vacuna contra la malaria humana? Stephen L. Hoffman y su equipo que trabaja en el programa de malaria del Instituto Naval de Investigación Médica de Bethesda escribieron recientemente en *Science* que una nueva vacuna, de la que forman parte dos proteínas de la malaria, había protegido de la infección a los múridos del ensayo. Queda por ver si tal éxito en los ratones se traduce en ayuda para la humanidad.

Hace un siglo, los colonos europeos de los trópicos evitaban los brotes de malaria con unas pocas dosis de agua de quinina. Los dos mil millones de personas que en la actualidad corren este riesgo no tienen un remedio tan fácil. La cloroquina, la mefloquina y otros medicamentos que se han usado para tratar la malaria están perdiendo su potencia a medida que cepas resistentes se desarrollan y se expanden por todo el mundo. Cada año se diagnostican alrededor de 110 millones de nuevos casos, con resultado fatal para uno o dos millones de personas.

Una vacuna eficaz contra la malaria podría prevenir muchas infecciones y reducir el problema de la evolución de la resistencia a los medicamentos. Pero el complejo ciclo biológico del parásito de la malaria dificulta la síntesis de una vacuna. Una picada de mosquito introduce esporozoítos, la primera forma del parásito, en un patrón humano. A los pocos minutos, los esporozoítos perforan las células hepáticas y empiezan a multiplicarse. En otras formas, destruyen los glóbulos rojos sanguíneos y transmiten la infección.

Con cada metamorfosis, señala Hoffman, las características inmunológicas del parásito cambian; significa ello que las vacunas apropiadas para una forma pueden fracasar ante otras. Con semejante estrategia, la malaria burla el sistema inmune. Aunque las personas que han sufrido la malaria en repetidas ocasiones parecen terminar desarrollando cierta resistencia parcial, nadie se hace nunca inmune de manera natural.

Han transcurrido más de veinte años desde que se demostró la posibilidad de evitar la malaria en roedores y seres humanos inyectándoles esporozoítos debilitados por la radiación. Pero no es tarea fácil producir esporozoítos en el laboratorio. Ni siquiera puede pensarse en conseguir los necesarios para proteger a los expuestos al parásito en Africa. Por eso se insiste en otra línea de investigación: descomponer el parásito para determinar los blancos y los mecanismos de la inmunidad inducida por el esporozoíto irradiado.

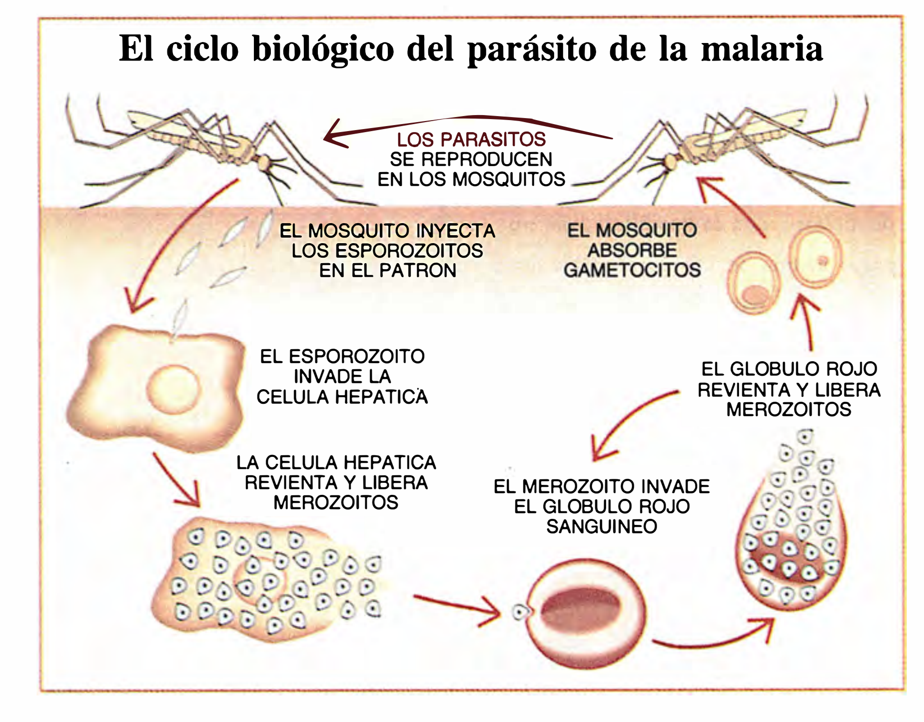
Buena parte de ese trabajo se concentró en la proteína del circumesporozoíto (CS), carácter de los esporozoítos altamente inmunogénico. Victor Nussenzweig, del Hospital Clínico de la Universidad de Nueva York, que ha trabajado en varias vacunas experimentales, asevera, sin embargo, que los productos basados en la proteína CS no han creado los niveles de anticuerpos esperados. En los animales modelo, el rendimiento es cabal; no así en humanos.

Hubo un segundo problema: las vacunas no estaban diseñadas para movilizar los linfocitos asesinos *T* (citotóxicos), glóbulos blancos de la sangre que son un potente componente del sistema inmune. Las células asesinas *T* activas son cruciales para la inmunidad ante la malaria.

El grupo de Hoffman tuvo más éxito porque no colocó todos sus huevos contra la enfermedad en un mismo cesto inmunológico. La vacuna del equipo presenta dos proteínas blanco: la proteína CS y una que se ha descubierto más recientemente, denominada proteína de la superficie del esporozoíto 2 (SSP2). El grupo insertó los genes para estas proteínas en células tumorales e inyectó en los ratones las células modificadas.

Muchos de los ratones así expuestos a una u otra de las dos proteínas quedaron protegidos frente a las infecciones provocadas por esporozoítos; cuando se les exponía a ambas proteínas conjuntamente, la protección abarcaba a todos los sujetos del ensayo, aun cuando se les sometiera a un número desorbitado de parásitos. Hoffman infiere de ello que la vacuna estimula a las células *T* para atacar a los parásitos en el hígado. Los anticuerpos producidos frente a las proteínas CS y SSP2 podrían, asimismo, matar a los esporozoítos.

Por más que los resultados son alentadores, los obstáculos para adaptar la vacuna eficaz en los múridos retrasarán muchos años la posibilidad de contar con una apta para el hombre. SSP2 es una proteína de un parásito de la malaria animal; todavía ha de encontrarse la proteína equi-



valente en los parásitos humanos. Puesto que inyectar células tumorales en personas es inaceptable, habrá que pensar en desarrollar otros métodos eficaces para obtener las proteínas.

Pájaros de mal agüero

Mientras unos se afanan por erradicar, con la obtención de una vacuna, la malaria producida por *Plasmodium falciparum*, la estirpe más mortífera, otros estudian su origen. Conclusión de los segundos: el hombre contrajo probablemente de las aves el parásito, en consecuencia indirecta de la introducción de la agricultura. De ser ello cierto, la enfermedad tendría menos de 10.000 años de edad.

Los parasitólogos se han dado cuenta de que *P. falciparum* es un tipo raro entre las cuatro especies de parásitos que causan las malarias humanas. Destaca por su virulencia extrema: hay que atribuirle más del 40 por ciento de todos los casos de malaria y más del 95 por ciento de todas las muertes por su infección. "Para un parásito, matar al patrón no suele ser una idea muy buena", explica Thomas F. McCutchan, del Instituto Nacional de la Alergia y Enfermedades Infecciosas (INAEI).

Esta observación sugiere que *P. falciparum* todavía no se ha adaptado del todo a su huésped humano. Otra diferencia es que los gametocitos, o formas sexuales, de los otros plasmodios humanos son redondos, mientras que el gametocito de *P. falciparum* tiene forma de hoz, como los de las malarias aviares.

Para comprobar el grado de parentesco entre *P. falciparum* y otros parásitos de la malaria, el grupo de McCutchan comparó las secuencias de ácidos nucleicos de las moléculas de ARN que se encuentran en los ribosomas de los parásitos. El número de diferencias en las subunidades de ARN ofrece pistas sobre el curso ramificador seguido por las especies.

McCutchan y sus colegas acaban de exponer sus resultados en los *Proceedings of the National Academy of Sciences*: *P. falciparum* guarda mayor similaridad con los plasmodios aviares que con los demás plasmodios humanos. En cambio, *P. vivax*, un plasmodio humano menos virulento, muestra una obvia relación con la enfermedad que se desarrolla en los primates y pudo haber coevolucionado con los seres humanos.

Según McCutchan, la malaria de *P. falciparum* sería una transferencia lateral a partir de aves. Hace mucho tiempo, un mosquito que había picado a un ave infectada de malaria pudo después haber picado e infectado a un ser humano. La mayoría de parásitos no hubieran sobrevivido; por desgracia, el antepasado de *P. falciparum* fue una excepción.

Dicha transferencia pudo haber tenido lugar en cualquier momento, pero McCutchan indica que la aparición de la agricultura resultó decisiva porque las sociedades agrarias permanecían estacionarias durante el tiempo suficiente para que la enfermedad se consolidara. El ciclo biológico del plasmodio no hubiera podido mantenerse con cazadores-recolectores, por la sencilla razón de que hubieran emigrado de la región.

McCutchan trae a colación las investigaciones de Mario Coluzzi, del Instituto de Parasitología de la Universidad de Roma, que demuestran que *Anopheles gambiae*, el mosquito que es el principal responsable de la transmisión de la malaria de *P. falciparum*, medra casi exclusivamente allí donde las personas viven o trabajan. Pequeñas charcas de aguas estancadas en tierras esbrozadas son lugares ideales de reproducción para el mosquito.

También es concebible, dice McCutchan, que la domesticación de las aves de granja desempeñara su papel en reunir aves infectadas y personas y en aumentar las oportunidades de transferencia lateral. Ulteriores estudios podrían aclarar qué patrón aviar dio albergue al antepasado más próximo de *P. falciparum*.

Imanes cuánticos

Si se borra de golpe el contenido del disco de su ordenador, no culpe a la máquina sino a la mecánica cuántica. De acuerdo con esa teoría, al enfriar un imán minúsculo hasta el cero absoluto, el polo norte y el polo sur intercambian su posición sin dificultad y, lo que es más, sin coste energético. Este efecto, denominado efecto túnel magnético-cuántico, podría afectar en menor grado a imanes a mayor temperatura, como los que forman un disco de almacenamiento de datos.

En honor de la verdad hay que considerar más probable que la electricidad estática borre la información de un disco que no un cambio de la polaridad de su base magnética. Sin embargo, algunos físicos empiezan a preocuparse por el papel que desempeña la mecánica cuántica en los pequeños imanes; en particular, a propósito de la nueva generación de sistemas de almacenamiento de datos. De 1950 a 1990 los átomos que se necesitaban para almacenar un bit de información bajaron de 10^{20} a 10^{9}. Si se prosigue a ese ritmo, para el año 2000 los dispositivos de almacenamiento de datos podrían traspasar la frontera de los 100.000 átomos por bit.

David D. Awschalom, del Centro de Investigación Thomas J. Watson de IBM, dirige un grupo que investiga el comportamiento de imanes de 100.000 átomos a temperaturas próximas al cero absoluto. Con las técnicas que se han desarrollado miden campos magnéticos con una sensibilidad un millón de veces superior a la de técnicas convencionales, como la resonancia magnética nuclear. Están convencidos de haber visto el efecto túnel y otros de la mecánica cuántica en tales sistemas magnéticos.

Su trabajo pone sobre el tapete cuestiones fundamentales del ámbito de la física y del dominio de la ingeniería. La física clásica gobierna el comportamiento de un imán formado por mil millones de átomos a temperatura ambiente; la mecánica cuántica, por contra, fija las reglas para un

Dispositivo de observación de imanes cuánticos

Los campos de los imanes pequeños pueden detectarse al inducir corrientes eléctricas en los hilos conductores negros. Las corrientes se amplifican mediante un dispositivo de interferencia cuántico y se miden.



INVESTIGACION Y
CIENCIA
Edición española de **SCIENTIFIC AMERICAN**

ha publicado sobre el tema, entre otros, los siguientes artículos:

Comunicaciones ópticas por cables submarinos, Raimundo Díaz de la Iglesia.
Número 117, junio 1986.

Nuevos metales, Bernard H. Kear.
Número 123, diciembre 1986

Nuevas cerámicas, H. Kent Bowen.
Número 123, diciembre 1986

Nuevos polímeros, Eric Baer.
Número 123, diciembre 1986

Materiales para el transporte terrestre, W. Dale Compton y Norman Louis A. Girifalco.
Número 123, diciembre 1986

Materiales para la navegación aerospacial, Morris A. Steinberg.
Número 123, diciembre 1986

Nuevos materiales y economía, Joel P. Clark y Merton C. Flemings.
Número 123, diciembre 1986

Materiales para la información y la comunicación, John S. Mayo.
Número 123, diciembre 1986

Electrones balísticos en semiconductores, Mordehai Heiblum y Lester F. Eastman.
Número 127, abril 1987

Transistores de arseniuro de galio, William R. Fresley.
Número 133, octubre 1987

Fractura del vidrio, Terry A. Michalske y Bruce C. Bunker.
Número 137, febrero 1988

Plásticos que conducen electricidad, Richard B. Kaner y Alan G. MacDiarmid.
Número 139, abril 1988

El dispositivo de efecto cuántico, ¿transistor del mañana?, Robert T. Bate.
Número 140, mayo 1988

imán de un átomo al cero absoluto. ¿Cómo guiarse cuando los imanes no se hallan en esas situaciones extremas?

El éxito de Awschalom tiene que ver con su destreza en integrar minúsculos imanes y formar un microsusceptómetro, ingenio que consta de dos bucles cuadrados de alambre, de 20 micrometros de lado. Se conectan de suerte tal que, al provocar una corriente eléctrica, el sentido de circulación en un cuadrado es opuesto al del otro. Los imanes se colocan en el interior de uno de esos bucles, donde generan una corriente tenue, que se amplifica con un dispositivo superconductor de interferencia cuántica (DSIC, O SQUID si abreviamos la expresión inglesa que lo designa)

Cada bucle se aloja en otro mayor. Al circular la corriente eléctrica por los bucles grandes, se generan dos campos magnéticos dotados de idéntica intensidad y orientación. Estos campos inducen la misma corriente por los dos bucles menores; ahora bien, por hallarse enfrentados, la corriente circula con sentido opuesto y se cancela. Por consiguiente, en los bucles interiores sólo habrá la corriente inducida por los imanes pequeños; bucles que únicamente sentirán el campo aplicado sobre esos imanes.

Para acomodar los imanes en los bucles, el grupo de IBM se sirvió de un microscopio de efecto túnel, cuya cámara se llenó con vapor de Fe(CO). Al crear una diferencia de potencial entre la punta del microscopio y la muestra, el vapor se "rompe" y se deposita sobre la muestra una gota de hierro y carbono formada por 100.000 minúsculos imanes o espines.

El grupo de IBM colocó un centenar de gotas en el interior de los bucles menores. (El trabajar con 100 gotas, y no con una sola, reporta la ventaja de percibir mejor su campo magnético combinado.) Las gotas se disponen a intervalos suficientes para que el campo magnético de una no influya en el comportamiento de sus vecinas. Si enfriamos los 100.000 espines de cada gota hasta 20 milikelvins, dos centésimas de grado por encima del cero absoluto, cada espín se entrelaza con su contiguo. La gota entera actúa a modo de un espín gigantesco, cuyo polo norte señala hacia arriba o hacia abajo.

La mecánica clásica postula que los espines no pueden cambiar su orientación sin haber recibido antes energía, limitación energética que impide el cambio de orientación de los espines. La mecánica cuántica establece, sin embargo, que hay posibilidades de que los espines rompan esa barrera, en un efecto parecido al que se da cuando un electrón atraviesa una barrera de energía.

Para someter a prueba sus teorías, el grupo de IBM comenzó por medir las propiedades magnéticas de las gotas sin campo magnético. Al aplicar la mecánica cuántica a este caso, resulta que los 100.000 espines de cada gota pasan, espontánea y frecuentemente, de una posición a otra. De ese modo los físicos introducen un campo magnético cuya polaridad cambia con el tiempo; si se ajusta la frecuencia del campo con la frecuencia natural de los imanes, entonces la oscilación se ve favorecida.

Tal fue el efecto resonante que observó el equipo de IBM. El experimento y algunos más aportaban así una prueba sólida de la existencia del efecto túnel magnético-cuántico. Constituía la primera observación real de la intervención de la mecánica cuántica en la realidad macroscópica, en sistemas magnéticos.

Pero, y Awschalom lo reconoce, teoría y experimento no acaban de encajar. La teoría indica la vía que lleva a inferir la frecuencia natural del efecto túnel a partir del tamaño de los imanes, su temperatura y la intensidad de los campos aplicados; ahora bien, realizados los cálculos, resulta un millón de veces mayor que el dato experimental medido por los investigadores de IBM. Ante semejante disparidad, responde Awschalom: "o estamos frente a un fenómeno completamente nuevo o bien la teoría es incompleta".

En cualquiere caso, él y su grupo deben concluir la difícil tarea de explicar el singular comportamiento de los imanes. Si le echan la culpa a la mecánica cuántica, quién sabe lo que guardan o pueden guardar en los imanes.

Exorcismo matemático

Pocos demonios azotan la ciencia, la estadística y el trabajo personal como el de resolver la incógnita *X*. A veces resulta trivial; pero otras no hay quien consiga integrar una función de múltiples variables. Cierta clase de esas integrales multivariantes ha resistido con impávida entereza el ataque de los matemáticos durante más de dos décadas.

Henryk Wozniakowski, matemático y experto en ordenadores, que si-

multanea sus clases en la Universidad de Columbia con su docencia en la de Varsovia, ha obtenido un algoritmo que rompe lo que, en la jerga matemática, se conoce por "maleficio de la dimensionalidad", al menos para la integración. Sus resultados, que despiertan el interés de físicos, químicos, biólogos e ingenieros, promete una resolución más rápida y eficaz de las funciones complejas.

El maleficio se refiere al incremento de la dificultad de un problema cuando el número de variables, o dimensiones, aumenta. Para una determinada precisión, el aumento de la dimensión en una unidad comporta multiplicar por 10.000 la dificultad de resolverlo. La obtención de una solución exacta para sistemas complejos, como las reacciones químicas, es imposible, ya que suelen implicar miles de dimensiones.

Los expertos hacen frente a esa aporía mediante aproximaciones a la solución real con métodos numéricos. El más conocido es el de Monte Carlo, según el cual se aproxima una solución de una integral evaluando el integrando en puntos elegidos aleatoriamente parar luego extrapolar. Los puntos necesarios para la selección elegida son los que cuantifican el coste del cálculo. Cuantos más puntos se eligen, más cercano del valor real se halla el resultado obtenido. Pero cuantos más puntos se escogen, más se alarga el tiempo de cálculo.

A pesar de todo, persisten funciones tan complejas que el coste de una aproximación es enorme, aun cuando se acepten errores moderados (representados por la épsilon griega, ϵ). Estos problemas complejos, sin embargo, serían más fáciles de resolver si se eligiera una evaluación promedio. Mientras que el método de Monte-Carlo garantiza la solución en un entorno de radio ϵ, mediante un algoritmo que da una evaluación promedio se obtienen soluciones con errores esperados o promedio pequeños. (La verdad es que el método de Monte Carlo elige la situación del peor de los casos; la respuesta cae dentro del error máximo.)

Aunque obtener soluciones con cierto error promedio parezca menos interesante que trabajar con una cota máxima para el error, las soluciones promedio no son necesariamente peores. Los matemáticos han demostrado que un algoritmo para este último tipo de soluciones garantiza un error al menos tan bueno como el error máximo; todo depende de los puntos de la muestra que se empleen para aproximar la solución. A diferencia del método de Monte Carlo, cuyos puntos son elegidos al azar, los algoritmos que dan una evaluación promedio recurren a puntos escogidos de una forma determinada. Pero no había forma de hallar tales puntos de la muestra; todos los esfuerzos se habían estrellado contra la dimensionalidad.

Wozniakowski se empeñó, pues, en formular un algoritmo eficaz de muestreo que aproximara la integral con el mínimo coste promedio, sin caer bajo el maleficio de la dimensionalidad. Dio con la clave cuando advirtió que el problema del promedio equivalía, matemáticamente hablando, al problema de la teoría de números conocido como discrepancia.

Para fortuna de todos, Klaus F. Roth, medalla Fields y adscrito al Colegio Imperial de Londres, resolvió el problema de la discrepancia en 1954. Wozniakowski descubrió la posibilidad de traducir el resultado de Roth al problema de la integración multivariante, y, con ello, seleccionar los puntos de la muestra que mengüen el error promedio.

El algoritmo es mucho más rápido que el método de Monte Carlo. El número de puntos necesarios para garantizar un error promedio menor que un error ϵ especificado de antemano es proporcional a $1/\epsilon$. En comparación, la cifra de puntos que hay que usar en el método de Monte Carlo es de $1/\epsilon^2$. (Este resultado parece implicar que, con el mismo número de puntos de la muestra, el método de Wozniakowski alcanza una precisión mayor que el de Monte Carlo. Algunos matemáticos no esconden, sin embargo, su reserva ante lo que es distinto criterio de error en uno y otro método: no podemos comparar alegremente cosas distintas.)

El trabajo de Wozniakowski apareció en el número de enero del *Bulletin of the American Mathematical Society*. Se trata de un trabajo bien trenzado que sabe plantear una cuestión nuclear de la teoría de la complejidad. El artículo ha despertado el interés de quienes laboran en otras disciplinas. Puede ser de gran ayuda en problemas de optimización, como el diseño de una cartera de negocios que optimice los beneficios. Los puntos de la muestra serían los índices de interés o los precios de las materias primas y de los productos básicos. Pero al teórico que es Wozniakowski, lo único que le movió a trabajar en el tema fue el reto de resolver un problema que estaba planteado desde hacía ya bastantes años.

Música electrónica con rostro humano

Se percibe, lejano, un suave rumor, de arena que golpea el pellejo de un tambor. Se sustituye, de repente, por un agudo tintineo, de martillos contra copas de cristal. Los golpes se aceleran, el tintineo sube. Mientras se escucha la pieza "Bug-Mudra" de Tod Machover, no es difícil advertir que buena parte de la misma ha salido de un ordenador. Pero si tenemos la oportunidad de ver a Machover dirigir la ejecución, la integración de los músicos humanos y electrónicos se nos queda grabada para siempre.

Enarbola en su derecha la batuta que orienta a los profesores; su izquierda, enfundada en un enorme "guante cibernético de datos" cubierto con sensores, dirige un ordenador Macintosh II que controla un grupo de sintetizadores. "Mi objetivo es construir cosas que confieran mayor creatividad al hacer musical, no sólo aparatos prostéticos", declara Machover, que es profesor en el Laboratorio de Medios del Instituto de Tecnología de Massachusetts.

Gran parte de la música electrónica continúa desarrollándose confinada a su lugar de nacimiento, los laboratorios de ciencias de la computación. Pero quienes han pasado años creando música por ordenador, Machover entre ellos, están decididos a incorporar sujetos humanos en la función. De ahí que trabajen en algoritmos complicados y en nuevos equipos que realcen la interpretación virtuosa de la música.

"Abundan los melómanos que no saben tocar ningún instrumento", resume Max V. Mathews, profesor del centro de investigación informática en música y acústica de la Universidad de Stanford. Hace veinte años, Mathews ayudó a encender la chispa de la revolución de la música sintética mediante algoritmos que ideó en los laboratorios Bell. Ahora espera volver a poner la música en manos de los oyentes mediante su "radiobatuta". Es, en muchos aspectos, una versión simplificada del guante de datos empleado por Machover.

La radiobatuta está provista de varios emisores, sencillos, que envían señales a una distribución de receptores; siguen éstos el movimiento de la batuta en tres dimensiones y comunican los datos a una orquesta simulada por ordenador. Un gesto hacia la izquierda da entrada a una línea de bajos; un movimiento hacia abajo acelera el *tempo*. "La gente no se limitará a cerrar los ojos y dejarse llevar por la música, antes bien podrá comprar las partituras y dirigir sus propias interpretaciones", predice Mathews.



HIPERVIOLONCELO de Tod Machover: híbrido de sensores e instrumento que improvisa a partir de la melodía de fondo. Foto de Peter Menzel.

Hay quien se propone conseguir "orquestinas musicales" formadas por ordenadores. "Queremos convertir el ordenador en un instrumento vivo", explica Miller Puckette, investigador del Instituto de Investigación y Coordinación Acústica/Música (IICAM) de París.

El grupo de Puckette ha ideado algoritmos que permiten a un ordenador seguir las notas que se ejecutan en un instrumento cercano y calcular su lugar en la partitura musical. "Hallar el tono constituye un problema formidable de análisis", confiesa. El IICAM está también elaborando algoritmos para lograr que el ordenador siga la línea de otros instrumentos. Tras percibir notas de una interpretación en curso, otros algoritmos establecidos generarán nuevas combinaciones —cuerdas o arpegios, por ejemplo— alrededor de ellas.

Otros investigadores se están especializando en "oyentes colaboradores"; se trata de programas informáticos, redes neurales incluidas, que siguen la música conforme se va tocando y son capaces de improvisar a partir de la estructura subyacente de la pieza. En ese frente labora David Wessel, director del centro de nueva música y audiotecnologías de la Universidad de California en Berkeley. "El colaborador informático toma fragmentos de lo que están tocando, trabaja sobre ellos y los reinyecta de manera inmediata en la ejecución", explica mientras da ejemplo improvisando sobre un tema de jazz que ejecuta con dos maestros y un oyente electrónico.

De manera análoga, los "hiperinstrumentos" (de Machover y otros) inventan nuevas interpretaciones de una ejecución en marcha. Los músicos tocan los hiperinstrumentos cual si fueran los tradicionales acústicos. Pero los híbridos electrónicos están cargados de sensores que captan cualquier inflexión de la interpretación. Los datos son la base de nuevos sonidos sintetizados que realzan la expresividad original del ejecutante.

Para conseguir que los ordenadores reciban la consideración de instrumentos ejecutores por derecho propio, se investiga también en la forma de "recapturar la intimidad del sonido", incorporando los matices de los instrumentos acústicos. Pase el arco por las cuerdas del violín, por ejemplo, y la nota producida será un rico conjunto de sonidos: el ataque inicial seguido por un crescendo que se va atenuando hasta un suave rumor. Los sintetizadores gobernados por ordenador han venido generando sonidos más uniformes.

Los sistemas convencionales de música suelen almacenar una colección de sonidos digitalizados —por ejemplo, todas la notas emitidas por un piano—, o un algoritmo específico

que dicta la combinación idónea de señales. Si se pide el *la* normal, el ordenador envía la descripción binaria de la nota a un sintetizador. Carla Scaletti, de Symbolic Sound, se propone sustituir la biblioteca de señales digitales por procesadores de señales digitales que el usuario pueda reprogramar; en particular, que pueda remodelar el *la* o crear un sonido nuevo.

Los sintetizadores basados en procesadores de señales digitales constituyen todavía un producto caro y de manejo difícil por parte de los músicos. Algunas compañías, sin embargo, se hallan interesadas en su desarrollo comercial. De cualquier forma, los avances en el procesado de señales digitales cambiará nuestro talante para con la música. Los maestros, por su lado, tendrán que convencer al auditorio del valor y la profundidad de los nuevos arrreglos orquestales.

Documentos tridimensionales

Examinar una casa que Antonio Gaudí proyectó pero nunca construyó resulta una óptima distracción visual con un ordenador dotado de capacidad gráfica. Se le dan al programa las órdenes pertinentes, y podemos comenzar a cambiar la iluminación para ver la casa al amanecer, en el resplandor del mediodía, con las vidrieras destellantes, o a la puesta del sol, cuando se alargan las sombras y se encienden la luces. Si preferimos contemplar la estructura desde otro ángulo o entrar en la casa, el ordenador le obedecerá. El problema es que no podemos llevarnos esas imágenes tridimensionales.

Ese sentido de impotencia se convirtió en reto para John R. Andrews, de Xerox, quien no paró de darle vueltas a las ventajas de poder llevar los documentos tridimensionales animados en una cartera; y no sólo los dibujos de un proyecto arquitectónico, sino cualquier asunto, desde el funcionamiento de una turbina hasta la salida de tinta de una impresora.

Con su equipo del Centro de Investigación Webster, Andrews ha encontrado una solución: un sistema para convertir los gráficos de ordenador en hologramas animados, imágenes tridimensionales que varían con el tiempo a medida que cambia el ángulo de visión del observador. Y habla ya del día en que las impresoras de los ordenadores de oficina producirán hologramas de gran formato.

Para producir los hologramas, Xerox codifica la luz procedente de la pantalla del ordenador convirtiéndola en las figuras de interferencia óptica que formarán un holograma cuando se las ilumine desde atrás. Cada figura se registra en una película en forma de una delgada banda horizontal. Una animación completa, como la de iluminar la casa de Gaudí desde la madrugada hasta el anochecer, puede contener hasta 70 bandas.

El conjunto de imágenes resultante se parece a una persiana. Un visor permitiría ver cada figura, pero como si abordáramos la escena a través de las rendijas de la persiana. Unas rendijas muy finas, porque los espacios se ajustan a la apertura media de la pupila del ojo, unos tres milímetros. Para que se pueda contemplar la escena entera, Xerox crea un segundo holograma a partir del primero. El proceso de copia hace las bandas invisibles al desenfocarlas respecto al ojo, como si la persiana se levantara por completo. El observador recibe información variable (del tiempo o el espacio) mediante el movimiento ascendente o descendente cabe la imagen iluminada.

En la complejidad del proceso recae el principal obstáculo para su comercialización. Lambertus Hesselink, físico de la Universidad de Stanford, sugiere usar un sistema híbrido para calcular y registrar hologramas digitales. En vez de calcular por puntos la totalidad de una pantalla, se divide en secciones, como un mosaico. Para cada trozo, el ordenador calcula las características físicas de la luz que se reflejaría de un objeto si existiese fuera del ordenador. Un haz láser suplanta la luz e ilumina esa zona.

Nadie se atreve a aventurar la fecha en que las impresoras empezarán a producir animaciones tridimensionales. Quizá se procederá con mayor lentitud de la deseada; por ejemplo, creando al comienzo sólo imágenes tridimensionales únicas.

FORTRAN redivivo

Década prodigiosa, dicen, la de los años sesenta. Si uno quería programar problemas de ciencia o ingeniería, había que usar FORTRAN. Implacable, desconocía la flexibilidad; pero no había nada más.

Vinieron, luego, otros lenguajes manejables, elegantes y potentes: C, BASIC, APL, LISP, PASCAL y otros. Pero los físicos y los ingenierons seguían empecinados en el FORTRAN. Analizaron los esfuerzos en puentes y aviones, modelaron el interior de los reactores de fusión y predijeron la evolución de las supernovas, todo ello en un lenguaje firmemente atado a las fichas perforadas.

Con el advenimiento de los multiprocesadores, que dividen los problemas en partes y resuelven cada una simultáneamente, los creadores de lenguajes comenzaron a cocinar un nuevo grupo de dialectos informáticos. Para esos procesadores paralelos inventaron SETL, NETL, ACTORS, LINDA, etcétera. Un grupo de ingenieros de programación acaba de terminar un trabajo en el que llevaban trece años empeñados: normalización de un lenguaje de programación más antiguo que, afirman, se adapta a múltiples tareas de programación paralela.

¿Cómo se llama? FORTRAN. Bueno, FORTRAN 90, también conocido por FORTRAN AMPLIADO. (Las versiones anteriores fueron la de FORTRAN 66, normalizado en 1966, y FORTRAN 77, que se aprobó en 1978.)

El nuevo FORTRAN se ideó con la mirada inicial puesta en la escritura fácil de programas científicos o de ingeniería. Pero gracias a un golpe de suerte, las mismas características que simplifican la vida de los programadores ofrecen también indicadores para pasar los programas en máquinas que incorporen cientos o incluso miles de procesadores, explica Dennis Hillis, creador del ordenador paralelo y fundador de Thinking Machines.

La principal innovación que incorpora el FORTRAN 90 reside en las instrucciones: pueden tratar matrices de datos —por ejemplo, la temperatura y velocidad del viento en los puntos de una reja que cubra el hemisferio occidental— como si constituyeran una sola unidad. El compilador (programa de ordenador que traduce las instrucciones FORTRAN a clave binaria) convierte esas instrucciones en órdenes que realizan operaciones sobre cada elemento de la matriz en paralelo. En el pasado, los programadores tenían que escribir instrucciones para abordar dato por dato.

George C. Fox, de la Universidad de Syracuse, estima que el FORTRAN 90 resulta ideal para la programación paralela de aproximadamente la mitad de todos los cálculos de ciencia e ingeniería que se ejecutan. Y agrega que cabe la posibilidad de refundir otros muchos programas para ajustarse a la clase de paralelismo que permite el FORTRAN 90.

La evolución no ha sido fácil. El trabajo de actualización del FORTRAN ha durado cerca de catorce años. Los cambios cuestan. Aunque las empresas jóvenes del ramo de la informática poseen ya compiladores que incorporan gran parte del FORTRAN 90, las más antiguas y establecidas han ido arrastrando los pies.

Se prevé una ampliación del FORTRAN 90 dentro de algunos años para aumentar el campo de programas paralelos fáciles de escribir e incluir nuevas propiedades. ¿Hasta cuándo seguirá remozándose lenguaje tan venerable?

Nubes estratosféricas polares y empobrecimiento en ozono

No suelen formarse nubes en la seca estratosfera antártica. Ahora bien, cuando lo hacen, conspiran químicamente con los halocarburos para crear el "agujero de ozono" que se abre todas las primaveras

Owen B. Toon y Richard P. Turco

Septiembre de 1987. Embarcaron en un DC-8 de la Administración Nacional de Aeronáutica y del Espacio (NASA). Formarían una treintena de expertos. El avión, que despegó de Punta Arenas, tomó rumbo sur. Al amanecer, ya en la península Antártica, se elevaron hacia la estratosfera, esa parte de la atmósfera que se extiende desde los 10 hasta los 50 kilómetros de altitud. A medida que el avión penetraba en el agujero de ozono, fenómeno hoy bien conocido, recibió la bienvenida de una nube imponente en forma de ojo, con un brillante iris rojo que rodeaba su pupila verde. Con el ER-2, un avión de gran altitud compañero, el DC-8 transportaba instrumentos para medir los aerosoles, gases y dinámica atmosférica de tales nubes, así como en la estratosfera circundante.

La misión tenía un objetivo claro: recoger datos que ayudaran a entender un fenómeno observado dos años antes, a saber, el de la correlación entre el empobrecimiento en ozono y la formación de las nubes estratosféricas.

A lo largo de los últimos cien años, los observadores desde el suelo habían venido registrando periódicamente la aparición de nubes estratosféricas sobre ambos polos, a alturas de unos 20 kilómetros. Las nubes se extienden a lo largo de 10 a 100 kilómetros en longitud y tienen varios kilómetros de espesor. Presentan una luminiscencia semejante a la de una madreperla; de aquí el nombre de nubes nacaradas, o madreperla, que se les da a veces.

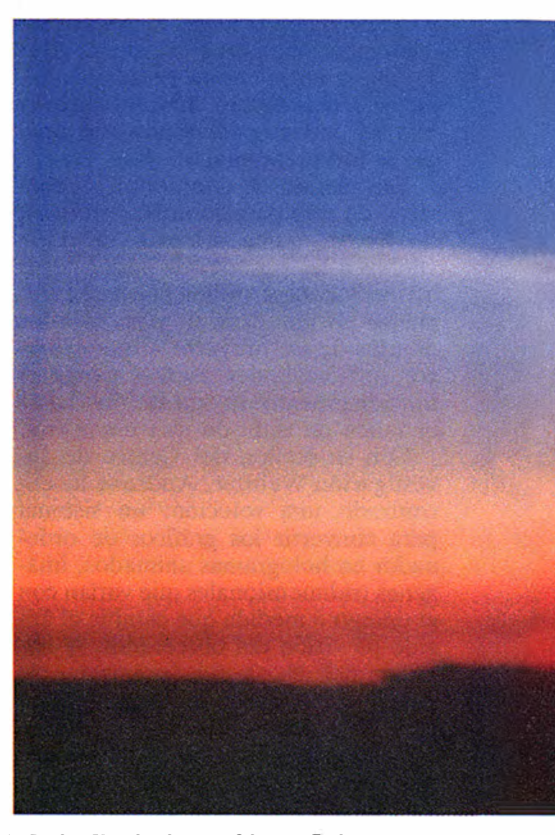
Además de las nacaradas, se han descubierto otras dos clases. La segunda no está formada por agua pura, sino por ácido nítrico. La composición química de la tercera es idéntica a la que caracteriza a las nubes nacaradas, pero se forma mediante un proceso del que resulta una nube mayor sin iridiscencia. Cuando esas tres clases de nubes se originan sobre los polos, hablamos de nubes estratosféricas polares (NEP, conocidas también con las siglas PSC por su denominación inglesa: "polar stratospheric clouds").

OWEN B. TOON y RICHARD P. TURCO aplican la información obtenida en exploraciones planetarias al estudio de la química atmosférica, radiación y microfísica. Toon, que recibió el doctorado en la Universidad Cornell, trabaja en el Centro de Investigación Ames de la Administración Nacional de Aeronáutica y del Espacio. Fue uno de los dirigentes de la misión aerotransportada a las regiones polares. Turco se doctoró por la Universidad de Illinois en Urbana-Champaign; da clases de física del aire en la Universidad de California en Los Angeles. Pertenecieron al grupo de trabajo que introdujo la noción de invierno nuclear.

La misteriosa belleza y exótica naturaleza de las NEP enmascaran su ominosa importancia. Recientes trabajos, incluida nuestra propia investigación, corroboran la idea de que esas nubes desencadenan el empobrecimiento en ozono de la estratosfera ártica. En la estratosfera antártica ayudan a crear el agujero de ozono.

En puridad no hay tal agujero. Se trata de una región que contiene concentraciones insólitamente bajas de ozono [véase "El agujero de ozono en la Antártida", por Richard S. Stolarski; INVESTIGACIÓN Y CIENCIA, marzo de 1988]. En el suelo, esta molécula, constituida por tres átomos de oxígeno, encierra un peligro para la salud; en la estratosfera, ese gas resulta decisivo para la vida en la Tierra. Aunque no importa ni siquiera una parte por millón de cuantos gases componen la atmósfera, él absorbe la mayor parte de la radiación ultravioleta que procede del Sol.

Tal radiación puede afectar el crecimiento y reproducción del fitoplancton, base de la cadena trófica marina. En lo concerniente al hombre, se establece una relación de causa a efecto entre la exposición excesiva a la radiación ultravioleta y los cánceres de piel, cataratas e inmu-



1. LAS NUBES ESTRATOSFERICAS se pueden formar sobre las regiones polares si el aire

nodeficiencias. Aunque situado en una región remota, el agujero de ozono constituye motivo de preocupación: las cambiantes configuraciones de la circulación transportan hacia el norte masas de aire empobrecidas en ozono. Ello puede presagiar una caída generalizada de ozono en toda la estratosfera.

En 1985, Joseph C. Farman y sus colaboradores, del Servicio Británico de Inspección Antártica, anunciaban que desde las postrimerías de los años setenta se venía observando un importante empobrecimiento en ozono en la estratosfera antártica. Las mediciones recabadas con el satélite *Nimbus 7*, y reelaboradas por Arlin Krueger, del Centro de Vuelos Espaciales Goddard, de la NASA, mostraron que, en el curso de los años, el empobrecimiento de primavera austral a primavera austral había ido empeorando, en general. Durante septiembre y octubre se pierde alrededor del 70 por ciento del ozono de la estratosfera antártica, que representa el 3 por ciento del ozono de la Tierra. Las mediciones realizadas por el grupo de David Hoffmann, de la Universidad de Wyoming, revelaron que el grueso de la caída ocurre entre 12 y 30 kilómetros de altura.

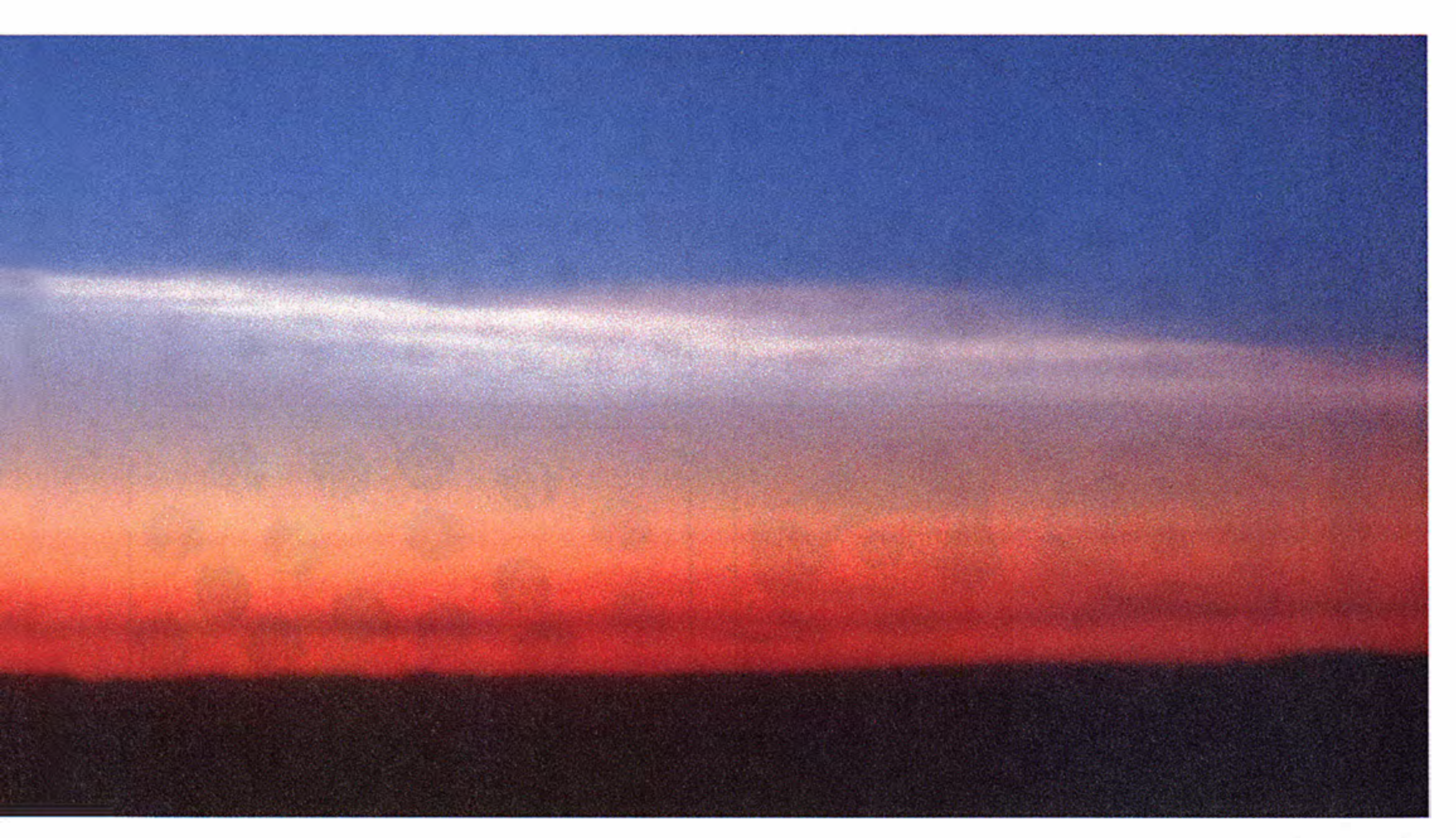
Se han propuesto distintas hipótesis para explicar la causa del agujero de ozono. Las expediciones se han encargado de aventar las erróneas. Para unos, los movimientos atmosféricos producían, por sí solos, el agujero de ozono; sostenían que las características de la circulación sobre los polos cambiaron quizá gradualmente, de suerte que ahora soplarían, llegada la primavera, vientos ascendentes sobre la Antártida. Esas corrientes sustituirían el aire estratosférico rico en ozono por aire pobre en el gas procedente de la troposfera (zona de la atmósfera por debajo de 10 kilómetros).

El equipo de Max Loewenstein, del Centro de Investigación Ames de la NASA, el grupo de Leroy E. Heidt, del Centro Nacional de Investigación Atmosférica (NCAR) de Boulder, y otros demostraron el error de esa línea argumentativa. Según los modelos dinámicos usados por los defensores de las hipótesis de la circulación, habría, a la altura del agujero de ozono, altas concentraciones de gases traza originados en el suelo. Sin embargo, las medidas de gases traza realizadas por distintos investigadores revelaron sólo bajos niveles, prueba de que el aire del agujero de ozono proviene efectivamente de grandes altitudes, donde abunda el ozono.

Una segunda clase de teorías atribuía la destrucción de nuestro gas a determinadas reacciones químicas. A tenor de una de las hipótesis madrugadoras, podría haber, en la vecindad del agujero de ozono, elevadas concentraciones de compuestos de nitrógeno reactivos; éstos son los principales agentes de la destrucción del ozono en la baja estratosfera. Se suponía que tal concentración debíase al efecto combinado de una mayor actividad solar y de la circulación atmosférica.

Proponíase que la mayor actividad solar producía formas reactivas de nitrógeno sobre el Polo Sur a grandes alturas. El movimiento de descenso del aire transportaría ese nitrógeno reactivo hacia la baja estratosfera, donde se había observado la pérdida de ozono. Pero Crofton B. Farmer y sus colegas, del Laboratorio de Propulsión a Chorro de la NASA, y el grupo de George H. Mount, del Laboratorio de Aeronomía de la Administración Nacional Oceánica y Atmosférica (NOAA), hallaron que las formas reactivas de nitrógeno también se habían empobrecido en el agujero de ozono. Quedaba así refutada la hipótesis.

El equipo de Farman ofreció una interpretación química sustitutoria, merecedora de una amplia aceptación. Se basa en un trabajo de mediados de los años setenta realizado por

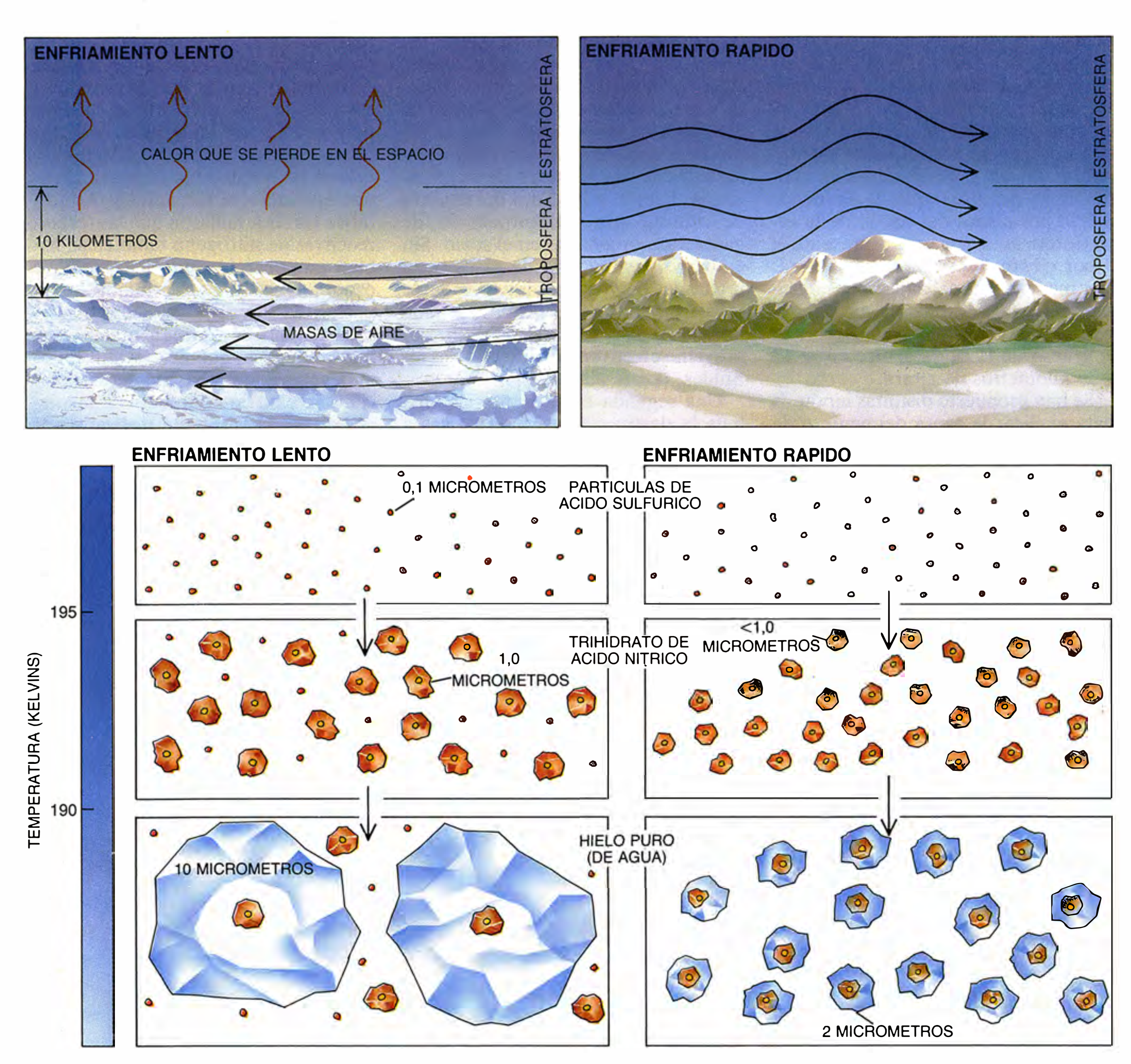


se enfría bastante durante el invierno. Las nubes de trihidrato de ácido nítrico se nos presentan en capas delgadas de color naranja oscuro. Las nubes de hielo puro aparecen blanquecinas. Nubes como éstas, desarrolladas sobre Stavanger, facilitan las reacciones químicas que destruyen el ozono.

Mario J. Molina, adscrito hoy al Instituto de Tecnología de Massachusetts, y F. Sherwood Rowland, de la Universidad de California en Irvine. Dicha teoría sugiere que los compuestos de cloro podrían ser los culpables del agujero de ozono. El cloro entra en la atmósfera a través, sobre todo, de los clorofluorcarburos (CFC) producidos por el hombre. Estos compuestos inertes se emplean en aplicaciones muy dispares: enfriadores para refrigeradores y aparatos de aire acondicionado, disolventes para la limpieza de circuitos impresos y agentes para la producción de materiales termoaisladores. Los CFC pueden sobrevivir de medio siglo a uno entero en la atmósfera. En muy pocos años, los vientos troposféricos distribuyen uniformemente las moléculas de CFC emitidas en un punto dado. En decenios alcanzarán la estratosfera media, a unos 30 kilómetros o más de altitud; allí, las rompe la radiación solar ultravioleta.

El cloro desprendido de las moléculas de CFC comienza su andadura en forma de cloro libre o reacciona con el ozono para constituir monóxido de cloro (ClO). Estas dos formas reaccionan ulteriormente para crear los depósitos de cloro, compuestos estables. Los integran, sobre todo, formas gaseosas del ácido clorhídrico (HCl) producido en la reacción del cloro libre con compuestos atmosféricos muy comunes: metano (CH_4) y nitrato de cloro ($ClONO_2$), originado éste de la reacción entre el óxido de cloro (ClO) y el dióxido de nitrógeno (NO_2).

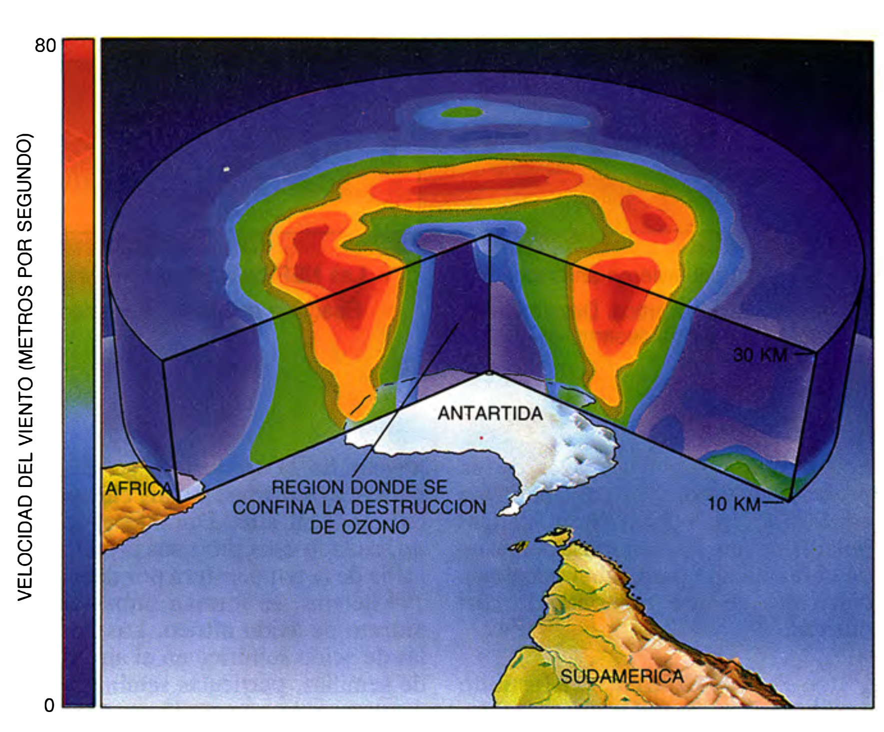
Los depósitos de cloro no destruyen por sí solos la capa de ozono. En estos compuestos, el cloro permanece



2. SE FORMAN NUBES ESTRATOSFERICAS POLARES cuando el aire se enfría hasta provocar la condensación del vapor. El enfriamiento puede producirse lentamente; así ocurre cuando la estratosfera radia calor al espacio o hay masas de aire que se deslizan por debajo y la obligan a elevarse (*parte superior izquierda*). El enfriamiento rápido se produce cuando el aire fluye sobre las montañas, creando unas ondas estacionarias que llegan hasta la estratosfera (*parte superior derecha*). Al caer la temperatura por debajo de 195 kelvins, las partículas de trihidrato de ácido nítrico se condensan alrededor de partículas de ácido sulfúrico en la estratosfera y forman nubes de trihidrato de ácido nítrico. Las nubes de hielo puro (de agua) se forman si la temperatura desciende por debajo de 190 kelvins; el vapor de agua se condensa sobre las partículas de trihidrato de ácido nítrico y de ácido sulfúrico. Bajo condiciones de enfriamiento rápido, una mayor fracción de tales partículas se convierte en núcleos de condensación.

inerte, sin poder reaccionar con el ozono. Los primeros modelos informáticos llegaron a la conclusión de que los CFC no tendrían por qué ejercer ningún efecto notable sobre la capa de ozono. Y añadían: sólo podrían destruirse pequeñas cantidades de ozono con la fracción de cloro que consiguiera escapar de los depósitos y volverse activa.

Resultaba obvia la existencia de algún mecanismo en la estratosfera antártica que liberaba bastante más cloro de los depósitos inertes. El grupo de Susan Solomon, del Laboratorio de Aeronomía de la NOAA, y el de Michael B. McElroy, de la Universidad de Harvard, aportaron las primeras pistas para su identificación. Sugirieron, en 1986, que la correlación observada entre el ciclo de destrucción de ozono y la presencia de nubes estratosféricas polares implicaba que, en las partículas de hielo de las nubes, se producían reacciones químicas que liberaban cloro de los depósitos.



3. VORTICE POLAR ANTARTICO, que confina la destrucción del ozono en el interior de un anillo de fuertes vientos, cuyas velocidades se representan mediante colores en el disco. La configuración geográfica no está dibujada a escala. El diagrama se basa en una imagen generada por ordenador hecha por Mark R. Schoeberl y Leslie R. Lait, del Centro de Vuelos Espaciales Goddard de la Administración Nacional de Aeronáutica y del Espacio (NASA).

Había una objeción palmaria contra la teoría de las nubes: éstas —suponíase— no solían darse en la estratosfera, con una humedad relativa, en promedio, del 1 por ciento. Además, allí el vapor de agua constituía sólo algunas partes por millón del aire, mil veces menos que en la troposfera donde se forman la mayoría de las nubes. Creíase, en efecto, hasta hace poco, que la única clase de nubes que existía en la estratosfera eran las madreperla. Estas se forman a altitudes entre 15 y 30 kilómetros; constituyen la versión estratosférica de las nubes lenticulares (en lente), que conocen muy bien quienes viven en regiones montañosas y ventosas. Las lenticulares se originan cuando el viento sopla a gran velocidad sobre las montañas. El viento crea una configuración estacionaria de las ondas de sotavento, que cursan ladera abajo de las montañas. En la porción ascendente de las ondas de montaña, el aire se dilata y enfría rápidamente. Si hay suficiente humedad, se condensará sobre las muchas partículas del aire. Las ondas se hacen así visibles en forma de nubes.

Si es estable la estratificación del aire, y el viento no cambia de dirección y velocidad a mayores alturas, la configuración de ondas estacionarias creada por las montañas se propagará hasta la estratosfera. Las nubes nacaradas tenderán entonces a formarse en las crestas de las ondas estacionarias. Lo harán por condensación sobre los aerosoles presentes.

El enfriamiento súbito y la condensación del vapor de agua forman las nubes nacaradas. Como las montañas crean ondas estacionarias, las nubes permanecen estacionarias, aun cuando el aire se mueva veloz a través de ellas. Los cristales de hielo capturan agua, al ser arrastrados hacia el interior de la nube por las corrientes de aire; crecen así hasta dos micrometros antes de haber acumulado todo el agua. Cuando los cristales alcanzan la porción descendente de las ondas estacionarias, el aire se comprime y se calienta, sublimando el hielo. Una sola nube puede extenderse entre 10 y 100 kilómetros de longitud. La energía de las ondas estacionarias puede bastar para alimentar una sucesión de esas nubes.

La distribución de tamaños en el interior del meteoro confiere a las nubes nacaradas su iridiscencia. Las partículas de hielo más chicas yacen en los bordes anterior y posterior; allí acaban de iniciar su desarrollo o están a punto de sublimarse. Las partículas mayores se forman en la parte central. Pero, en cualquier posición donde se encuentren, las partículas difractarán la luz solar incidente. La intensidad de la luz difractada depende de la longitud de onda de la luz incidente y del tamaño de la partícula. Debido a ello, cuando observamos las nubes bajo ángulos no muy grandes respecto al Sol, aparecen brillantes. Los colores siguen el contorno de la nube, reproduciendo la distribución de tamaños de las partículas.

Las nubes nacaradas indicaban a los meteorólogos que la estratosfera estaba lo bastante fría para que se formaran cristales de hielo, al menos en las regiones polares. De su extremada sequedad se infiere que la temperatura debe estar por debajo de 190 kelvins (−83 grados Celsius). Sólo en el invierno antártico se mantienen tales temperaturas durante suficiente tiempo. (También existen nubes en la estratosfera ártica, pero se forman menos frecuentemente porque las temperaturas medias de invierno son allí más altas que en la Antártida.)

Los datos de los satélites revelaron lo que los observadores de tierra no podían ver. El instrumento de medida de aerosoles estratosféricos (MAE II, o SAM II de acuerdo con la expresión inglesa que abrevia: "Stratospheric Aerosol Measurement"), lanzado a bordo del satélite *Nimbus 7* en 1978 y gestionado por M. Patrick McCormick, del Centro de Investigación Langley de la NASA, detectó partículas en el aire examinando la luz rasante en el borde de la Tierra. El MAE II mostró que las nubes estratosféricas existían sobre la Antártida, aun cuando las temperaturas bajaran hasta sólo 195 kelvins (−78 grados Celsius). Tales temperaturas eran demasiado altas para que se formaran nubes nacaradas. Cabía, pues, suponer que las



- Comienzo del invierno antártico.
- Formación del vórtice. Desciende la temperatura y se desarrollan nubes.
- Las NEP desnitrifican y deshidratan la estratosfera por medio de la precipitación.
- Reacción del ácido clorhídrico y el nitrato de cloro en partículas nubosas: cloro libre.
- Temperatura invernal en su punto más bajo.

nubes se originaban a través de un proceso distinto. Respaldaban esa conclusión las nubes estratosféricas polares recién descubiertas, demasiado desarrolladas para haber nacido de corrientes de aire tramontando cordilleras.

Con Paul J. Crutzen, del Instituto Max Planck de Química en Maguncia, y Frank Arnold, del Instituto Max Planck de Física en Heidelberg, sugerimos en 1986 que estas nubes debían diferir en composición de las nacaradas, formadas de agua pura condensada sobre partículas suspendidas. Las teorías químicas relativas a la destrucción del ozono exigen la eliminación del nitrógeno reactivo; sin ello, éste aprisionaría al cloro en nitrato de cloro, uno de los principales depósitos de tal elemento. Las nubes, dedujimos, podrían servir de sumidero de nitrógeno. Si así ocurría, tenían que estar constituidas por una forma congelada del ácido nítrico (HNO_3) con tres moléculas de agua por una molécula de nítrico. El trihidrato de ácido nítrico ($HNO_3 \cdot 3H_2O$), ése es el compuesto, no sólo explica la eliminación del nitrógeno, sino que además se condensa a temperaturas más altas que el agua pura.

En observaciones independientes, varios grupos dirigidos por David W. Fahey, del Laboratorio de Aeronomía de la NOAA, Bruce W. Gandrud, del NCAR, y Rudolph F. Pueschel y Stefan A. Kinne, del Centro de Investigación Ames de la NASA, confirmaron nuestras teorías. Determinamos también, con otros, que esas nubes no se formaban, en general, por brusco enfriamiento del aire obligado a elevarse por las montañas. Antes bien, las produce un proceso de enfriamiento lento. La estratosfera polar en invierno radia energía al espacio durante la larga noche polar; con el tiempo, regiones extensas registran las temperaturas que posibilitan la formación de partículas de nube. Añádase que los sistemas meteorológicos de la troposfera se deslizan bajo la estratosfera polar, elevando y enfriando el aire. Durante el invierno, cuando esos procesos provocan la caída de la temperatura por debajo de 195 kelvins, se forman nubes de trihidrato de ácido nítrico. Las partículas de ácido sulfúrico en el aire sirven de semillas, partículas seminales que provienen de gases sulfurosos liberados en procesos biológicos naturales y antropógenos. La circulación atmosférica transporta las partículas emitidas en el suelo hasta la baja estratosfera. También las erupciones volcánicas explosivas pueden lanzar gases sulfurosos directamente a la estratosfera. Las partículas, de un tamaño aproximado de 0,1 micrometros, pueden alcanzar una concentración considerable años después. En 1982, El Chichón, situado cerca de Pichucalco (México), expulsó unos cinco millones de toneladas de azufre a la estratosfera.

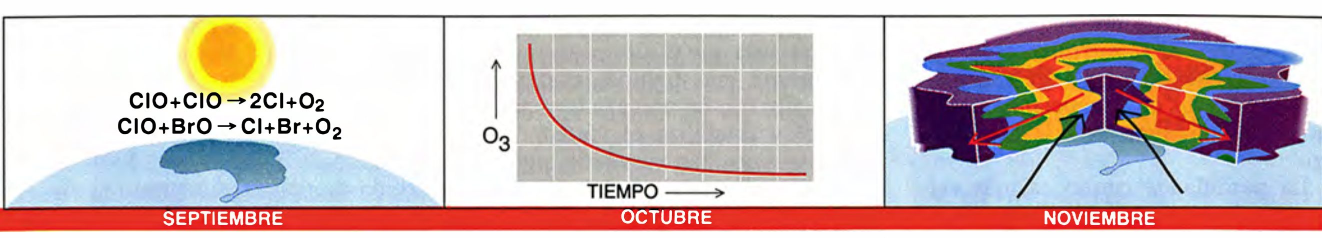
El enfriamiento lento puede producir nubes estratosféricas que se extiendan en un radio muy amplio. Los equipos de McCormick y Edward B. Bowell, del Centro de Investigación Langley de la NASA, emplearon radares de láser (lidares) aerotransportados para cartografiar las nubes una por una. Descubrieron que éstas solían constar de capas múltiples de kilómetros de espesor y se desplegaban a lo largo de varios millares de kilómetros. Comparadas con las nubes nacaradas, las de trihidrato de ácido nítrico poseen menos masa y son más tenues, por cuya razón cuesta observarlas a simple vista.

Además de las nacaradas y las de trihidrato de ácido nítrico, puede formarse una tercera clase de nubes estratosféricas, lo que acontece si en el invierno antártico la temperatura cae por debajo de 190 kelvins. Cuando el aire se enfría, el vapor de agua se condensa sobre partículas suspendidas, desarrollando nubes de hielo puro (de agua). Las partículas seminales son las de ácido nítrico ("germinadas", a su vez, sobre otras de ácido sulfúrico); constituyen las nubes de trihidrato de ácido nítrico.

Lo mismo que las nubes nacaradas, este tipo de NEP contiene hielo puro. Aunque suelen clasificarse juntas las dos clases de nubes de hielo puro, se las distingue en razón de su velocidad de formación (las nacaradas se crean por enfriamiento rápido). Estas nubes de hielo puro no son tan corrientes como las de trihidrato de ácido nítrico, especialmente en el Artico, a causa de las bajísimas temperaturas necesarias para su formación.

A diferencia de las nubes de hielo de agua de enfriamiento rápido, las de hielo de agua de enfriamiento lento suelen escapar a la observación en tierra. En ambas clases de nubes, la masa de agua condensada viene a ser la misma, si bien las partículas de las nubes de enfriamiento lento son mayores que las de las nubes nacaradas. El rápido enfriamiento que insta el desarrollo de las nubes nacaradas transforma casi todos los aerosoles suspendidos en núcleos de condensación; el enfriamiento lento usa sólo una pequeña fracción de las partículas presentes. En virtud de ello, las nubes nacaradas contienen un gran número de cristalitos de hielo, de un tamaño aproximado de dos micrometros. Las nubes de enfriamiento lento contienen menos cristales, aunque mayores, con un tamaño que excede los 10 micrometros. Las nubes de hielo puro de enfriamiento lento alojan menos partículas por unidad de volumen y no reflejan la luz con la perfección de sus primas nacaradas.

Las tres clases de NEP —de trihidrato de ácido nítrico, de hielo puro de enfriamiento lento y de hielo puro de enfriamiento rápido (nacaradas)— ejercen un papel clave en el empobrecimiento en ozono de la Antártida. Las NEP pueden activar el cloro en sus superficies y consumir el nitrógeno reactivo, impidiéndole que transfiera el cloro a sus depósitos.



SEPTIEMBRE

- Primavera austral: la luz solar retorna al centro del vórtice y las NEP desaparecen al subir la temperatura.
- Los ciclos catalíticos ClO-ClO y ClO-BrO destruyen el ozono.

OCTUBRE

- Se alcanzan los niveles más bajos de ozono.

NOVIEMBRE

- Desplome del vórtice polar estratosférico.
- El aire rico en ozono, de latitudes medias, repone el ozono de la estratosfera antártica.
- El aire pobre en ozono se extiende sobre el Hemisferio Sur.

Además, las nubes de enfriamiento lento de agua pura y las de trihidrato de ácido nítrico pueden arrebatarle a la estratosfera todo el nitrógeno.

Los estudios de laboratorio de Molina y Ming-Taun Leu, del Laboratorio de Propulsión a Chorro, y de Margaret Tolbert, del Instituto Internacional de Investigación de Stanford, mostraron que la reacción entre el ácido clorhídrico y el nitrato de cloro —los dos componentes que mantienen el cloro en su estado inactivo— se produce sobre superficies de hielo puro y de trihidrato de ácido nítrico. Esa reacción, que rinde cloro molecular (Cl_2) y ácido nítrico, es tan lenta que carece de importancia sin la presencia de partículas sólidas.

Bajo la luz solar de la primavera antártica, el cloro molecular se disocia presto en cloro atómico altamente reactivo, precipitando el ciclo catalítico ClO-ClO. En este ciclo, el átomo de cloro, recién liberado, rompe la molécula de ozono para dar oxígeno molecular y monóxido de cloro. Molina descubrió que el monóxido de cloro en fase gaseosa reacciona consigo mismo, formando su dímero (Cl_2O_2). La luz solar disocia fácilmente el dímero en átomos libres de cloro, lo que conduce a una ulterior destrucción de ozono. El cloro mantiene así un papel catalítico en el empobrecimiento de ozono.

Si hubiera dióxido de nitrógeno, se combinaría rápidamente con monóxido de cloro para aprisionar el cloro en la molécula depósito inerte, nitrato de cloro; se suspendería así el ciclo catalítico ClO-ClO. Pero las NEP impiden la reacción porque convierten cualquier nitrógeno presente en ácido nítrico.

El grupo de James G. Anderson, de la Universidad de Harvard, y Robert L. deZafra y Philip Solomon, de la Universidad estatal de Nueva York en Stony Brook, hallaron niveles sorprendentemente altos de ClO en el agujero de ozono antártico: unas 500 veces la cantidad acostumbrada en latitudes medias a la misma altura. Ante tal concentración, el ciclo catalítico ClO-ClO da cuenta, en su mayor parte, de la caída de ozono observada. Un solo átomo de cloro se basta para destruir millares de moléculas de ozono antes de encontrar compuestos reactivos de hidrógeno o nitrógeno que lo devuelvan de nuevo a sus depósitos. Prosigue la investigación de laboratorio de todas esas reacciones; se busca con ello acotar los caminos que toma la química y determinar con mayor precisión las velocidades de reacción.

El cloro activo liberado por las NEP participa de manera destacada en otro importante proceso catalítico en que interviene el bromo. Libérase éste a la atmósfera en distintas actividades humanas, siendo componente notable de los extintores. La reacción podría responsabilizarse de hasta el 20 por ciento de la destrucción de ozono.

Debemos a McElroy y sus colaboradores la descripción del proceso reactivo. El bromo extrae un átomo de oxígeno del ozono y forma monóxido de bromo (BrO), compuesto que reacciona con el monóxido de cloro. La reacción produce oxígeno molecular y libera los átomos de bromo y cloro, que de nuevo reaccionan con el ozono, repitiendo el proceso. La corroboración de la realidad de tal proceso catalítico vino con las observaciones realizadas en 1987 por William H. Brune, de la Universidad estatal de Pennsylvania, y Anderson, de Harvard; éstos midieron niveles apreciables de óxido de bromo en el agujero de ozono antártico.

Además de producir reacciones químicas que convierten las formas inertes de cloro en formas reactivas (y las reactivas de nitrógeno en inertes), las nubes estratosféricas polares desnitrifican, es decir, eliminan nitrógeno de la estratosfera. Las nubes de hielo puro de enfriamiento lento podrían constituir los agentes principales de la desnitrificación. Las partículas de hielo puro de las nubes, amén de desarrollarse sobre núcleos de ácido nítrico, absorben ácido nítrico en forma de vapor. Las partículas, que alcanzan hasta 10 o más micrometros, caen desde la estratosfera antártica en forma de nieve; y en ese proceso la estratosfera se desnitrifica y deshidrata al mismo tiempo. No parece que las nubes nacaradas sean capaces de eliminar nitrógeno de la atmósfera. Las rápidas corrientes de aire que las atraviesan tienden a sublimar las partículas de hielo puro antes de que se inicie la precipitación.

No sólo las nubes de hielo puro de enfriamiento lento; también las de trihidrato de ácido nítrico se manifiestan capacitadas para desnitrificar. Sus partículas miden característicamente alrededor de un micrometro de diámetro, magnitud limitada que se explica por el escaso ácido nítrico que la estratosfera contiene y tamaño que le permite mantenerse en suspensión. Algunas nubes se forman tan despacio que sus partículas crecen por encima de un micrometro y pueden así caer de la estratosfera. Esa descripción del proceso se ha visto respaldada con observaciones de la estratosfera ártica, que se ha desnitrificado pero no deshidratado. Además, las mediciones acometidas por Browell con lidar muestran que algunas nubes de trihidrato de ácido nítrico contienen partículas de más de un micrometro.

La actual teoría NEP-clorofluorcarburos de la formación del agujero de ozono explica muchas observaciones. La liberación antropogénica de CFC, principalmente en el Hemisferio Norte, es responsable del empobrecimiento en ozono del Hemisferio Sur: la dilatada vida media de los halocarburos en la atmósfera facilita su distribución uniforme en la misma. El agujero de ozono se produce en la vecindad de la Antártida durante la primavera porque su formación requiere

la presencia de nubes estratosféricas polares, que se desarrollan sólo durante la época más fría del año. Los primeros rayos de sol inician las reacciones químicas que destruyen el ozono.

La pérdida de ozono es más evidente sobre la Antártida que sobre el ártico. Por la sencilla razón de que la estratosfera antártica es más fría, lo que posibilita la formación de un número mayor de nubes, en particular por debajo de 20 kilómetros. Un número mayor de esos meteoros significa producir más átomos de cloro reactivo y eliminar más compuestos de nitrógeno, lo que redunda en una mayor pérdida de ozono.

Una diferencia tal vez más importante entre ambos polos se refiere a la longevidad del vórtice polar, anillo de fuertes vientos que confina la destrucción de ozono. En la Antártida, el vórtice permanece intacto durante todo el invierno polar, e incluso hasta mitad de primavera. La destrucción del ozono comienza en septiembre con el retorno de la luz solar. La pérdida de ozono alcanza su máximo en octubre. En el Artico, donde la configuración de la circulación difiere bastante de la que presenta la Antártida, el vórtice se desintegra mucho antes de que llegue la primavera polar (marzo y abril).

Hay un mecanismo de realimentación que puede alargar aún más la vida del vórtice antártico. El ozono absorbe luz solar y calienta la atmósfera; la caída de la concentración del gas determina que el aire permanezca frío durante más tiempo; esa situación de aire frío favorece la formación de NEP y estabiliza el vórtice. Las mediciones nos han enseñado que, en los últimos diez años, la temperatura en el interior del vórtice ha disminuido y, aumentado, el tiempo durante el cual el vórtice permanece intacto. El grupo de Lamont R. Poole, del Centro de Investigación Langley de la NASA, ha demostrado que las NEP acostumbran formarse cuando desciende la temperatura, completando así la consolidación del proceso.

El temprano derrumbamiento del vórtice ártico hace difícil estimar la magnitud de la pérdida de ozono en el Hemisferio Norte. Ahora bien, por las mediciones que Brune y Anderson realizaron en 1989 sabemos que, en ambos vórtices polares, se producen casi las mismas cantidades de cloro a alturas de unos 18 kilómetros. Al parecer, existen en el Artico condiciones que favorecen la formación de un agujero de ozono.

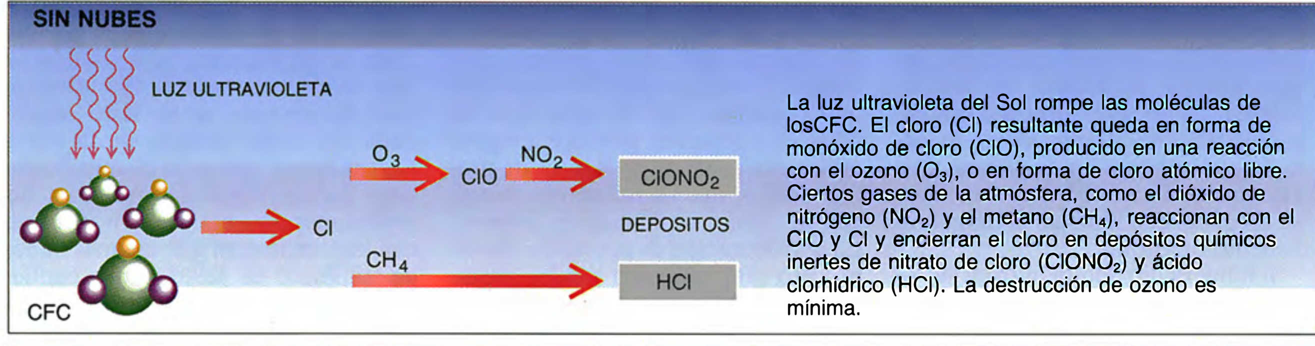
Por su parte, Browell y Michael H. Proffitt, del Laboratorio de Aeronomía de la NOAA, descubrieron, en 1989 también, la pérdida de ozono experimentada en extensas regiones de la estratosfera ártica por encima de 18 kilómetros: la caída se cifraba en torno al 6 por ciento de la cantidad total de ozono sobre el Artico. En comparación, la pérdida antártica total de ese gas supone, en promedio, un 50 por ciento o más. No se produjo una merma mayor porque las regiones inferiores a los 18 kilómetros poseían una temperatura tal que impedía el desarrollo de NEP. En el futuro, la pérdida de ozono aumentará conforme vayan creciendo los niveles atmosféricos de cloro.

A menos que la temperatura ártica sufra un descenso notable, el empobrecimiento en ozono del Artico nunca rivalizará con la pérdida en la Antártida. Una temperatura media invernal más baja posibilitaría la formación de NEP sobre una región de mayor altura, nubes que podrían, además, persistir durante más tiempo.

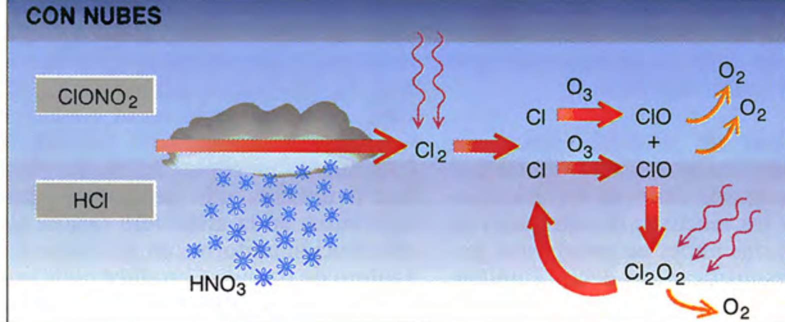
La creación del agujero de ozono requiere, sabemos, la formación de nubes estratosféricas polares y un vórtice polar estable; por ello, aquél queda necesariamente confinado cerca de los polos, donde vive muy poca gente. Lo que no signfica, en absoluto, que la pérdida de ozono se limite a las regiones polares.

El derrumbe del vórtice antártico insta la dispersión, por el Hemisferio Sur, de bolsas de aire pobre en ozono. En diciembre de 1987, Rodger Atkinson, de la Oficina Australiana de Meteorología, informó sobre los niveles de ozono más bajos que se conocían sobre el sur de Australia y Nueva Zelanda tras el desplome del vórtice polar. Por culpa de esa dispersión, las bolsas de aire pobre en

Las nubes estratosféricas polares y la destrucción del ozono por el cloro



La luz ultravioleta del Sol rompe las moléculas de losCFC. El cloro (Cl) resultante queda en forma de monóxido de cloro (ClO), producido en una reacción con el ozono (O_3), o en forma de cloro atómico libre. Ciertos gases de la atmósfera, como el dióxido de nitrógeno (NO_2) y el metano (CH_4), reaccionan con el ClO y Cl y encierran el cloro en depósitos químicos inertes de nitrato de cloro ($ClONO_2$) y ácido clorhídrico (HCl). La destrucción de ozono es mínima.



Las nubes estratosférica polares inician reacciones químicas que liberan cloro molecular (Cl_2) de los depósitos. El ciclo catalítico ClO-ClO comienza una vez que la luz solar rompe la molécula de Cl_2. Los átomos de cloro, Cl, reaccionan con el ozono, para formar ClO y oxígeno (O_2). El ClO se dimeriza (Cl_2O_2), molécula que la luz solar disgrega en Cl y O_2. El cloro entonces vuelve a atacar al ozono. Las NEP impiden que se formen depósitos eliminando el nitrógeno de la atmósfera mediante la precipitación de ácido nítrico (HNO_3). El bromo también destruye ozono por reacción catalítica con el ClO (no representada).

ozono provocan una pequeña pérdida media de ozono en todo el hemisferio. A mayor abundamiento, Adrian F. Tuck, del Laboratorio de Aeronomía de la NOAA, ha sugerido que el vórtice antártico podría comportarse como un procesador químico: el aire rico en ozono penetraría en el vórtice y el pobre en ozono se sacaría fuera, hacia el resto del hemisferio. Una expedición con un avión de gran altitud, planeada para 1993, pondrá a prueba esta hipótesis y aportará datos más precisos sobre el empobrecimiento en ozono de la Antártida.

No toda la química destructora del ozono está confinada en las regiones polares. Las partículas de ácido sulfúrico que se encuentran en la estratosfera desencadenan también la reducción del ozono liberando cloro de sus depósitos moleculares. Y lo que encierra mayor interés: esas partículas convierten nitrógeno reactivo en formas inertes e impiden así la formación de depósitos de cloro. La pérdida de ozono causada por los aerosoles de ácido sulfúrico, sin embargo, es menos importante que la causada por las NEP, debido a la menor masa y tamaño de partícula.

Hoffmann y Solomon han observado que la nube volcánica creada por la erupción de El Chichón mermó bastante los niveles de ozono. La cantidad de aerosoles en forma de sulfatos desprendida rivalizó con la masa de NEP de trihidrato de ácido nítrico. Si se produjera una gran erupción volcánica a principios del próximo siglo (cuando los niveles de cloro sean elevados), asistiríamos a una apreciable pérdida global de ozono.

Aunque quedan muchos detalles por conocer, sí sabemos ya la naturaleza del eslabón fundamental entre las partículas estratosféricas y la destrucción del ozono. A pesar de lo cual, ni el refinamiento ulterior de la química estratosférica parece que pudiera ofrecernos un medio técnico eficaz para detener ese empobrecimiento. Reponer el ozono que se pierde cada año —cuantía equivalente a la masa total de la población humana— requeriría, por lo menos, el uno por ciento de la producción energética de los EE.UU., y ello sin contar la tarea formidable de elevar el gas hasta el vórtice antártico.

No se ve, pues, otra solución que apelar a un acuerdo internacional para limitar la producción de CFC. En 1987, los EE.UU. y otros países industrializados acordaron restringirla con la firma del Protocolo de Montreal "sobre substancias que empobrecen la capa de ozono". El protocolo acordaba en un comienzo reducir las emisiones, antes del año 2000, a un 50 por ciento de los niveles de 1986. Pero con esa merma la concentración de cloro, a finales del siglo que viene, habría doblado la actual.

Conforme los científicos iban conociendo mejor la química de la destrucción del ozono y del papel que en ella desempeña el cloro, las naciones cobraron conciencia de la necesidad de tomar medidas más drásticas. En 1990 se convino en suspender por completo la producción de halocarburos para este cambio de siglo.

Los niveles atmosféricos de cloro continuarán, sin embargo, aumentando durante las próximas décadas. Quedan grandes cantidades de CFC en los refrigeradores, acondicionadores de aire y espumas termoaislantes; en buena parte, y andando el tiempo, pasarán a la atmósfera. La sustitución general de los CFC tardará, como mínimo, diez años en empezar a lograrse.

Se prevé que la cantidad de cloro atmosférico alcanzará su máximo durante el primer decenio del próximo siglo. Los halocarburos se distinguen por su larga vida media, razón por la cual quizás el cloro no retorne a sus niveles anteriores a la producción del agujero de ozono hasta mediados del siglo próximo, si no más tarde.

Por consiguiente, la destrucción del ozono se agravará con los años, durante algunos decenios, hasta el punto presumible de que el agujero llegue a duplicar su superficie antártica. Se ha apostado por la hipótesis de acuerdo con la cual cierta pérdida mayor de ozono distorsionaría menos los ecosistemas y las actividades humanas que el control y eliminación inmediata de los CFC. Una apuesta sin suficientes garantías.

BIBLIOGRAFIA COMPLEMENTARIA

PROGRESS TOWARDS A QUANTITATIVE UNDERSTANDING OF ANTARCTIC OZONE DEPLETION. Susan Solomon en *Nature*, vol. 347, n.º 6291, págs. 347-354; 27 de septiembre de 1990.

THE DYNAMICS OF THE STRATOSPHERIC POLAR VORTEX AND ITS RELATION TO SPRINGTIME OZONE DEPLETIONS. Mark R. Schoeberl y Dennis L. Hartmann en *Science*, vol. 251, págs. 46-52; 4 de enero de 1991.

FREE RADICALS WITHIN THE ANTARCTIC VORTEX: THE ROLE OF CFCs IN ANTARCTIC OZONE LOSS. J. G. Anderson, D. W. Toohey y W. H. Brune en *Science*, vol. 251, págs. 39-46; 4 de enero de 1991.

Factores de transcripción y control de la expresión génica

Estas proteínas interaccionan con áreas específicas del ADN y controlan la expresión de los genes. El estudio de estos factores y su modo de acción nos permitirá entender el desarrollo del cáncer, la inmunidad y el envejecimiento

Antonio Celada

El material genético es el conjunto de caracteres que se heredan de los progenitores y determina las diferencias entre especies o individuos. En definitiva, hace que seamos como somos, individualmente distintos. El material genético que posee un organismo proviene de ambos progenitores a partes iguales. En el proceso de la fecundación intervienen dos células germinales, cada una de las cuales posee la mitad de material genético que una célula somática. Tras fusionarse el espermatozoide con el óvulo, todas las células que se producen a partir de aquí (con excepción de las progenitoras) tendrán la misma cantidad de material genético. En el curso del desarrollo, las células se multiplican y diferencian, es decir, adquieren características propias.

En el organismo humano existen unos 200 tipos de células que llevan a cabo funciones biológicas muy determinadas. Estas diferentes células, aunque tienen muchas funciones en común, poseen otras que les son exclusivas. Por sólo nombrar unas pocas: los linfocitos se ocupan de la defensa del organismo, los miocitos son las células del músculo y las neuronas lo son del sistema nervioso. Esto quiere decir que, de los 100.000 genes que contiene cada célula humana, una serie de ellos se expresan en todas las células y otros en determinado tipo celular, lo que hace que la diversificación de las células sea sólo el resultado de la diferente expresión de una serie de genes. En cada célula del organismo, los genes representan una especie de diccionario que contiene todas las palabras. Sin embargo, para hacer una frase hay que saber escoger y utilizar las palabras apropiadas y esto es lo que ocurre precisamente en el transcurso del desarrollo.

ANTONIO CELADA trabaja en la Fundación de Investigación del Cáncer de La Jolla, desde 1985. Pertenece también al Consejo Superior de Investigaciones Científicas. Tras licenciarse en medicina por la Universidad de Valladolid en 1972, se trasladó a Ginebra, donde residió hasta 1980. En 1974 se doctoró por la Universidad de Ginebra y en 1977 por la de Salamanca. Pasó dos años en la Universidad de Washington y otros tres en la Fundación de Investigación de la Clínica Scripps. Su campo de interés se centra en el mecanismo de activación de los macrófagos mediado por el interferón gamma, para lo cual ha llevado a cabo estudios clínicos, de biología celular y de biología molecular en el área de los factores de transcripción, tema del artículo.

A finales del siglo XIX se descubrió que los portadores de la información genética eran los cromosomas, estructuras nucleares que se hacen visibles en el núcleo al comienzo de la división celular. En los cromosomas, el material genético lo constituye el ácido desoxirribonucleico (ADN). Estas moléculas están formadas por dos cadenas cuyos eslabones están constituidos por cuatro tipos de nucleótidos, adenina (A), citosina (C), guanina (G) y timidina (T), unidos a un esqueleto de carbohidratos fosforilados. Las dos cadenas están enrolladas entre sí en doble hélice, a modo de collar de perlas de doble vuelta en el que las cuentas serían los nucleótidos. En cada cadena, los nucleótidos se encuentran en la parte interna y son complementarios entre sí, es decir, que los de una cadena se unen específicamente a los de la otra, A con T y C con G. Previamente al proceso de división celular, las dos hebras de la molécula de ADN se separan, para convertirse en moldes donde se engarzarán los nucleótidos complementarios; así se origina una nueva doble cadena de ADN y se perpetúa el material genético.

La combinación de los nucleótidos por grupos de tres en tres codifica los distintos aminoácidos cuyas combinaciones de secuencias determinarán, a su vez, las diferentes proteínas, que son la base estructural del organismo. El ADN se puede comparar a una cinta magnética que, al ser leída, establece la síntesis del ácido ribonucleico (ARN) a través de los nucleótidos complementarios; es decir, A en ADN será T en ARN, C será G, G será C y T será el uracilo (U), nucleótido que distingue al ARN del ADN.

Entre la síntesis de ADN y la producción de proteínas hay una serie de procesos intermediarios sobre los que el organismo puede ejercer un control y modificar la producción proteica. Estas etapas son las siguientes: 1) transcripción del gen, iniciación de la producción de ARN mensajero (ARNm, el ARN que viajará del núcleo al citoplasma, donde se llevará a cabo la síntesis proteica); 2) proceso de corte y empalme ("splicing"), es decir, según los casos, se transcribirán ciertas áreas de la molécula de ADN pero no otras, produciendo un número elevado de combinaciones; 3) transporte del ARNm del núcleo al citoplasma; 4) estabilización del ARNm en el citoplasma, para evitar que se degrade, en cuyo caso no sería reconocido en los ribosomas; 5) traducción del ARN en los ribosomas, esto es, síntesis de aminoácidos y, por consiguiente, de proteínas; 6) control de la actividad proteica mediante activación, inactivación o compartimentalización.

Dentro de cada molécula de ADN existe un punto en el que se inicia la transcripción. Esta se lleva a cabo mediante la activación de una enzi-

ma, la ARN polimerasa II. La parte inicial de la molécula de ADN contiene los elementos reguladores, es decir, segmentos específicos que controlan el que la transcripción se inicie. Responden al nombre de promotores. Existen otras regiones cuya misión consiste en regular el nivel de expresión, una vez iniciada la transcripción y que se denominan amplificadores. Contrariamente a los promotores, que se encuentran siempre por delante del lugar donde se inicia la transcripción (unos 200-300 nucleótidos antes), los amplificadores pueden alojarse muy distantes, lo mismo en la parte anterior que en la posterior de la molécula de ADN. En los últimos años se ha demostrado que una serie de proteínas entran en interacción específica con los promotores o los amplificadores: las que conocemos por factores de transcripción.

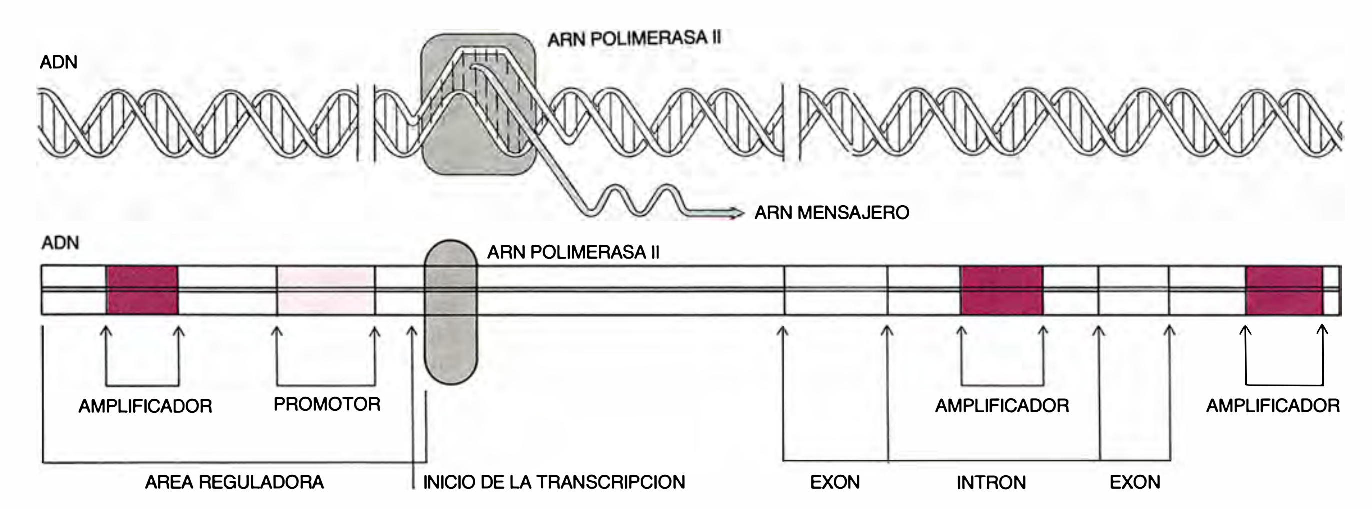
¿Cómo identificar los factores de transcripción? Hay que empezar por delimitar las áreas del ADN que controlan la expresión del gen. Para ello se provocan deleciones (eliminación de partes) o mutaciones (intercambios de nucleótidos) del ADN de la región reguladora, y se mide el efecto que producen sobre la transcripción. Así se localizan, con exactitud, el promotor o los genes multiplicadores ("amplificadores"). La cuantificación del efecto de las modificaciones inducidas por las deleciones o mutaciones se puede llevar a cabo por diversos métodos. A tenor de uno de los más utilizados, se une la parte del ADN que se sospecha contiene los elementos reguladores del gen que nos interesa a otro gen cuya actividad se pueda medir fácilmente, verbigracia, el cloranfenicol-acetil-transferasa (CAT). Si el ADN resultante de la transfección (introducción en las células de ese ensamblaje) contiene los elementos reguladores, expresará la enzima CAT. La actividad de esta enzima se puede medir fácilmente mediante la conversión de cloranfenicol radiomarcado en acetil o diacetil cloranfenicol; la cantidad de estos productos es directamente proporcional a la cantidad de enzima presente. Otro sistema consiste en transferir ("transfectar") el gen en el que estamos interesados, y del que hemos eliminado una de las partes que se expresan; ello origina un ARNm de menor tamaño, fácilmente distinguible del endógeno producido por la propia célula. Por último, existen sistemas de transcripción *in vitro* en los que se añaden extractos provenientes de núcleos (proteínas obtenidas del núcleo celular).

Una vez localizadas las regiones de interés, tales como promotores o multiplicadores, se procede a identificar las proteínas que se unen específicamente a esas zonas. Uno de los métodos más utilizados es el llamado "retraso del ADN en geles". En este caso, se radiomarca el fragmento de ADN que contiene la región reguladora y se incuba con extractos obtenidos a partir de núcleos celulares. Se lleva a cabo una electroforesis en gel de poliacrilamida (matriz que tiene la propiedad de separar ADN o proteínas según su tamaño mediante el paso de una corriente eléctrica continua).

Si en la mezcla de extractos nucleares con ADN existe una proteína que se fija a éste, se formará un complejo de proteína/ADN cuya migración en el gel será más lenta, en comparación con el ADN libre; en la autorradiografía, el complejo aparecerá como una banda en la parte superior en tanto que al ADN libre corresponderá otra banda, ahora en la parte inferior.

La especificidad de la reacción entre las proteínas y el ADN se estudia mediante experimentos de competición. Si la reacción es específica, la banda de retraso desaparece al añadir el mismo fragmento de ADN sobrante sin marcar radiactivamente (unas 20-100 veces más), pero persiste si se añaden fragmentos sobrantes de ADN sin marcar que corresponden a otras secuencias de ADN con las que no guardan relación. Los geles de retraso de ADN sirven no sólo para identificar proteínas que se fijan al ADN, sino también para estudiar las características de la interacción proteína/ADN, tales como la constante de equilibrio o la cinética de la reacción.

La localización exacta del lugar de fijación de la proteína al ADN se determina mediante el método de "improntas". En este caso, el ADN, radiomarcado en un extremo, es tratado en presencia de extractos nucleares con dimetilsulfato, agente que metila las guaninas (les añade un grupo CH_3) de una forma aleatoria, es decir, una sola G por molécula de ADN. Si el factor de transcripción se une a una G, la protegerá la acción del dimetilsulfato. Tras la migración en geles, el ADN libre o en forma de complejo con las proteínas se aísla. Se le trata entonces con piperidina. Este agente fracciona las moléculas de



1. DENTRO DE CADA MOLECULA DE ADN existe un punto en el que se inicia la transcripción (producción de ARNm). Se lleva ésta a cabo mediante la activación de la enzima ARN polimerasa II. La parte anterior de la molécula contiene los elementos reguladores, es decir, segmentos específicos de ADN que controlan el arranque de la transcripción: son los promotores. Otros segmentos tienen por misión la de regular el nivel de expresión, una vez iniciada la transcripción: son los multiplicadores ("amplificadores"). En oposición a los promotores, que se encuentran siempre por delante del lugar donde se inicia la transcripción (unos 200-300 nucleótidos antes), los amplificadores pueden localizarse, lo mismo en la parte anterior que en la posterior de la molécula de ADN. Se ha demostrado que varias proteínas interaccionan con los promotores o los amplificadores: son los factores de transcripción. Llamamos intrones a los segmentos de ADN que no se transcriben y, exones, a los que se transcriben.

ADN donde las G están metiladas; es decir, corta en un solo lugar cada molécula de ADN creando un número igual de fragmentos de muy diversos tamaños.

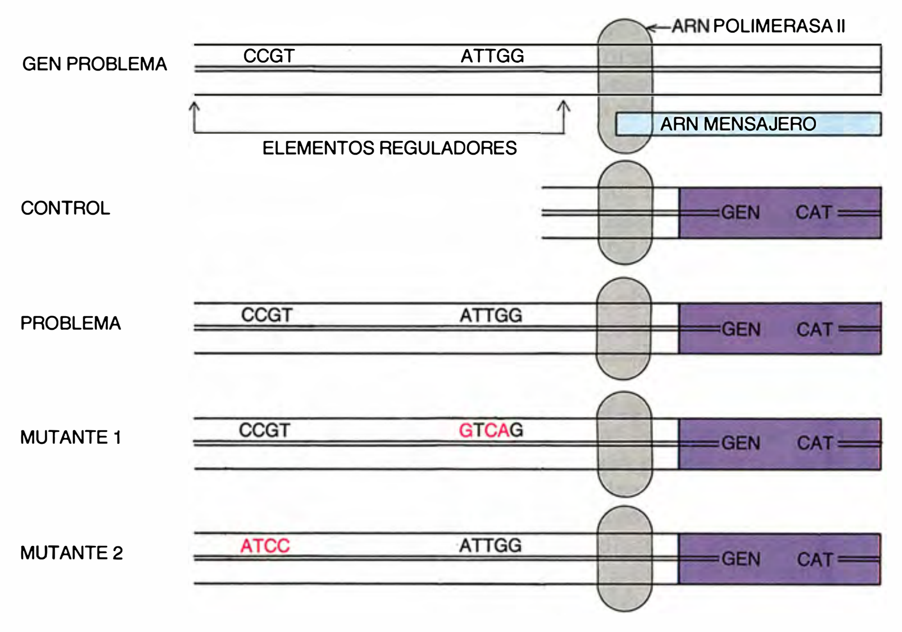
Al promover la migración del ADN así tratado en un gel de secuenciación (de un espesor de 1 milímetro o menos), aparecerán múltiples bandas que habrán corrido de acuerdo con su tamaño. Si en los extractos añadidos al ADN hay una o varias proteínas que se fijan con suficiente afinidad (intensidad), lo protegerán de la metilación y posterior digestión, con lo cual el ADN no se fraccionará. En la posición del gel correspondiente al fragmento de ADN protegido de la digestión por la proteína, no veremos ya bandas, salvo la de control que indica la posición exacta de contacto entre la proteína y el ADN.

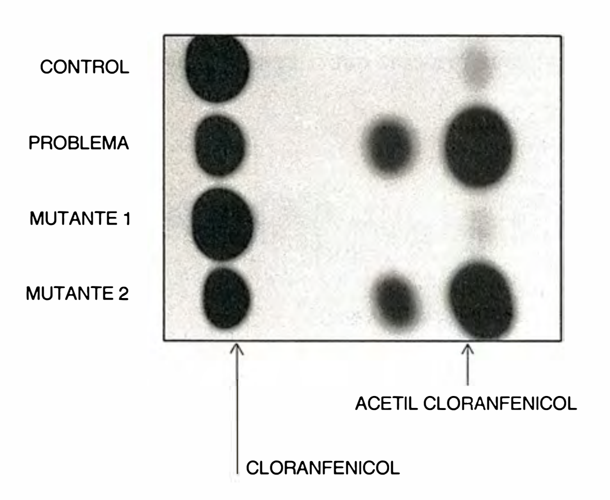
El siguiente paso consiste en clonar, es decir, aislar el gen que codifica la proteína reguladora que se fija al ADN para averiguar su secuencia de nucleótidos. Aquí se empieza por purificar la proteína a partir de grandes cantidades de extracto celular. Se recurre, para ello, a los métodos cromatográficos, que sacan partido de las propiedades físico-químicas de la proteína, tales como el peso molecular o la carga, propiedades que las distinguen del resto de las proteínas de los extractos nucleares. La proteína suele obtenerse totalmente pura tras el paso a través de una columna de oligonucleótidos sintéticos (fragmentos de ADN), que contienen la secuencia del ADN a la que se une la proteína, lo que se habrá determinado previamente mediante el método de "improntas".

Una vez purificada la proteína, conoceremos la secuencia de aminoácidos que la componen. Prepararemos luego oligonucleótidos sintéticos, que se emplearán de sonda para el clonaje del ADN. La sonda radiomarcada se hibrida (dado que contiene secuencias complementarias) con el ADN de una genoteca (conjunto del ADN total de la célula). De esta forma se puede identificar el ADN que determina la proteína deseada. Una vez aislado el ADN del resto del que contiene la genoteca, se puede introducir en un vector de expresión y, posteriormente, ya sea en bacterias o mediante sistemas *in vitro*, producir la proteína recombinante. Así podemos controlar que el ADN clonado determina la proteína (factor de transcripción) que se fija al ADN.

Por otro lado, la proteína sintética puede servir para estudios de transcripción *in vitro*, o estudios de asociación con otros factores de transcripción y para producir anticuerpos. La obtención de anticuerpos reviste particular interés por cuanto permite determinar la localización intracelular o el estado de fosforilación o glicosilación de la proteína natural. También se puede estudiar la actividad del factor de transcripción transfiriéndolo ("transfectándolo") a células que no lo expresan. Ello se lleva a cabo introduciendo al mismo tiempo el promotor donde se supone que actúa el factor de transcripción con un indicador como el CAT, del que hemos hablado antes.

En factores de transcripción que son específicos de tejido, se puede introducir el ADN que codifica para el factor de transcripción en embriones; en virtud de lo cual se expresará en todos los tipos celulares, lo que nos ayudará a comprender su actividad funcional. Los animales transgénicos resultantes de esos embriones pueden generarse de tres formas distintas: mediante microinyección del ADN en huevos fecundados, por infección de embriones con vectores retrovíricos y mediante transferencia de genes en células madre ("stem"), embrionarias.

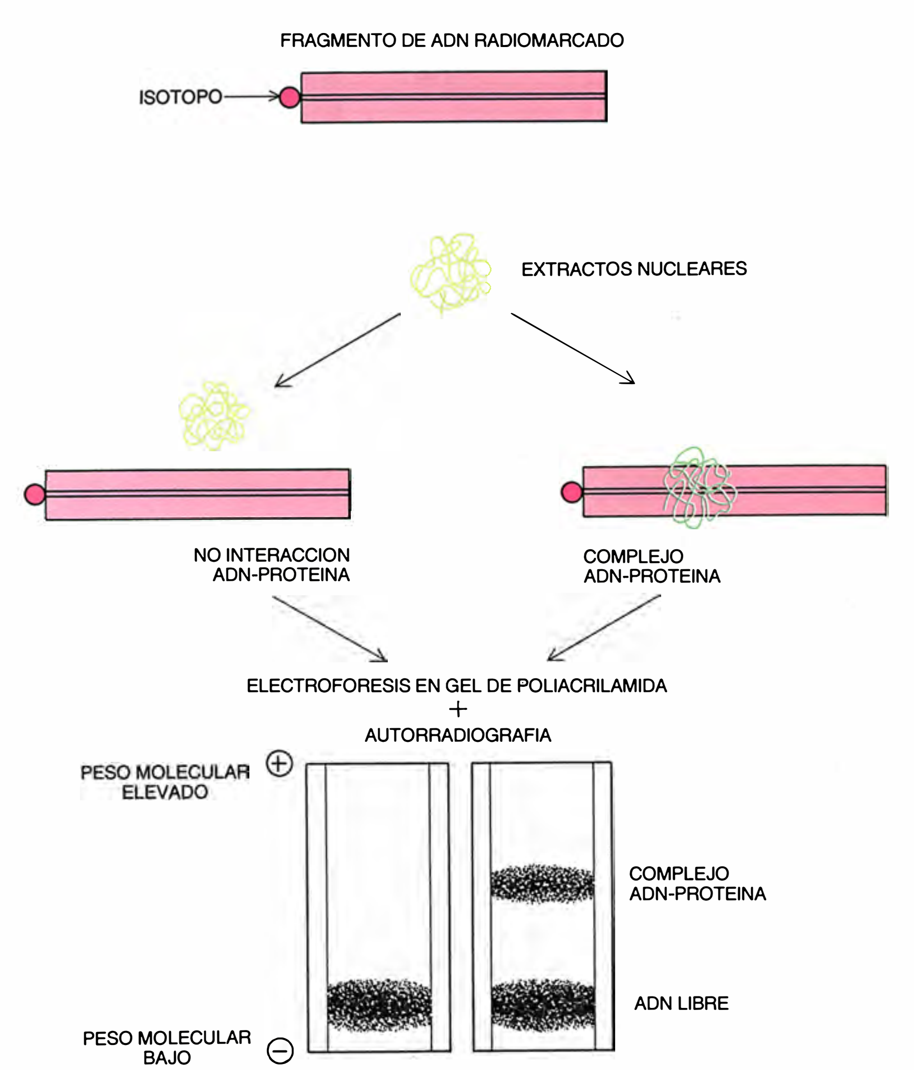




2. IDENTIFICACION DE LOS FACTORES DE TRANSCRIPCION y su secuencia de pasos. Se empieza por reconocer qué segmentos del ADN controlan la expresión del gen. Para ello se provocan deleciones (eliminación de partes) o mutaciones (intercambios de nucleótidos) del ADN en la región reguladora (*señaladas en la figura con un asterisco*), midiendo el efecto que producen sobre la transcripción. Se localiza así con exactitud el promotor o los multiplicadores ("amplificadores"). Hay varios métodos de cuantificar las modificaciones inducidas por las deleciones o mutaciones. Uno de los más utilizados consiste en unir la parte del ADN que se sospecha contiene los elementos reguladores del gen que nos interesa a otro gen cuya actividad se mide fácilmente, como el cloranfenicol-acetil-transferasa (CAT). El ADN resultante de la transfección (introducción en las células de este constructo), si contiene los elementos reguladores, expresará la enzima CAT. La actividad de esta enzima se mide a través de la conversión de cloranfenicol radiomarcado en acetil o diacetil cloranfenicol; la cantidad de estos productos es directamente proporcional a la cantidad de enzima presente.

Contamos con un método alternativo para la clonación de los factores de transcripción. Conocida la región del ADN que interacciona con el factor, se construye un oligonucleótido sintético, que se utiliza de sonda en una genoteca de expresión, es decir, una vez que el ADN de la genoteca ha expresado su mensaje produciendo proteínas. Cuando la sonda se una a una proteína de la genoteca, se hallará fácilmente el ADN que la determina. Pero se trata de un método que está más expuesto a errores y sólo puede usarse para factores de transcripción que tengan un solo componente (algunos de estos factores necesitan dos componentes que interaccionan entre sí para obtener una fijación eficaz al ADN).

La clonación de una serie de factores de transcripción, operada a lo largo de los últimos 5 o 6 años, nos ha ofrecido la oportunidad de investigar los pormenores de su interacción con el ADN. Al clonar el gen que expresa el factor de transcripción, podemos provocar deleciones o mutaciones (intercambio de nucleótidos) y delimitar así qué parte se une al ADN. Y, por otro lado, disponemos de los resultados de los estudios estructurales acometidos con microscopía electrónica, análisis cristalográfico con rayos X o resonancia magnética nuclear. Gracias a esos trabajos conocemos elementos estructurales que son comunes a diversos grupos de factores de transcripción.

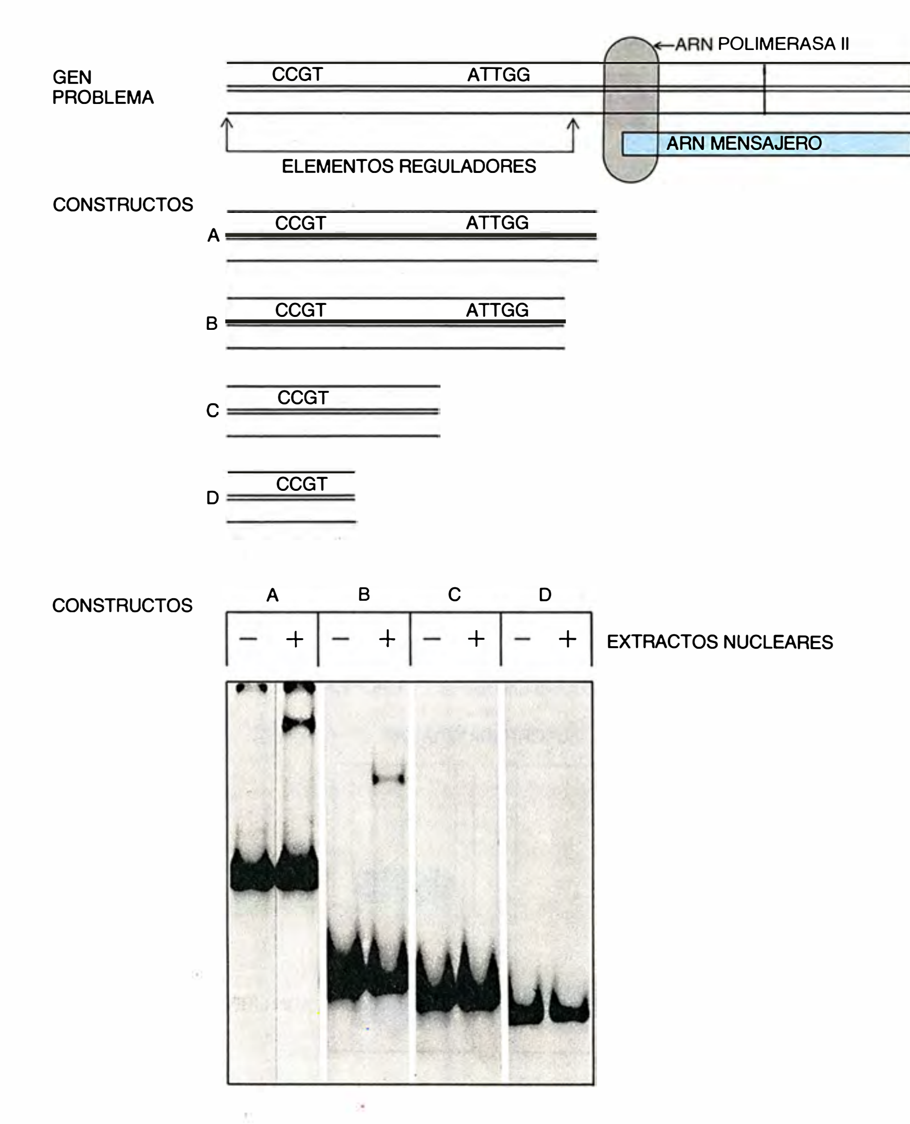


3. LOCALIZADAS LAS AREAS DE INTERES, tales como promotores o multiplicadores ("amplificadores"), deben identificarse las proteínas que se unen específicamente a dichas áreas. Se recurre sobre todo al método del "retraso del ADN en geles". Se marca radiactivamente el fragmento de ADN que contiene la región reguladora y se incuba con extractos obtenidos a partir de núcleos celulares. Luego, se lleva a cabo una electroforesis en gel de poliacrilamida (matriz que tiene la propiedad de separar ADN o proteínas según su tamaño mediante el paso de una corriente eléctrica continua). Si en la mezcla de extractos nucleares con ADN existe una proteína que se fija al ADN, se formará un complejo proteína/ADN, cuya migración en el gel será más lenta en comparación con el ADN libre; en la autorradiografía aparecerá en forma de banda en la parte superior, en tanto que el ADN libre formará otra banda, pero en la parte inferior.

La primera estructura que se estudió, y la mejor caracterizada hasta el momento, es la llamada hélice-giro-hélice. A pesar de la considerable variabilidad de la secuencia de aminoácidos de los distintos factores de transcripción, el motivo hélice-giro-hélice no altera apenas su geometría. Los aminoácidos de la "hélice de reconocimiento" se hallan en contacto directo con las bases del ADN. La otra hélice se coloca alrededor y se une al ADN de forma no específica. El motivo hélice-giro-hélice se identificó primero en una familia de proteínas de *Drosophila*, que se han mantenido a lo largo de la evolución y que intervienen controlando la diferenciación embrionaria, aunque se desconoce su mecanismo preciso de acción. Estas proteínas poseen un dominio (región estructural o funcional) de unos 60 aminoácidos, que se mantiene en distintas especies; responde al nombre de homeodominio ("homeobox"). En este dominio es donde se halla el motivo hélice-giro-hélice. Se ha comprobado la capacidad de las homeoproteínas para fijar ADN; ello, sumado a otros datos, abona la idea de que constituyen factores de transcripción.

Uno de los primeros factores de transcripción en clonarse fue el TFIIIA. Se trata del modelo en "dedos de zinc" que se halla presente en algunos factores de transcripción. La región de este factor que se une al ADN alberga nueve series repetidas de 30 aminoácidos y cada una contiene dos cisteínas y dos histidinas. Entre estas dos parejas se inserta un átomo de zinc y los aminoácidos entre ellas se proyectan como un dedo cuya punta contacta directamente con el ADN. Este motivo, "dedo de zinc", es necesario para mantener la estructura de la proteína, aunque no interaccione directamente con el ADN.

Otro motivo presente en algunos factores de transcripción es la llamada "cremallera de leucinas": cada siete aminoácidos hay un residuo de leucina, en una estructura helicoidal que permite que las leucinas se encuentren cada dos giros en el mismo lado de la hélice. Las cremalleras de leucina no se unen directamente al ADN, sino que facilitan la dimerización (unión de proteínas) mediante la interdigitación de las dos hélices que contienen leucina. La dimerización produce una estructura proteica que se une al ADN a través de una región adyacente. Existen otros factores de transcripción cuya estructura no se adapta a ninguno de los tres modelos mencionados, siendo muchos de ellos en el momento actual objeto de estudio.



4. ESPECIFICIDAD DE LA REACCION entre las proteínas y el ADN. Para demostrarla, se realizan experimentos de competición entre fragmentos radiactivamente marcados y fragmentos de ADN sin marcar que porten la secuencia correspondiente al promotor o amplificador que hemos identificado previamente o con secuencias de ADN distinto. Los geles de retraso de ADN ayudan a conocer los pormenores de la interacción proteína/ADN; por ejemplo, la constante de equilibrio o la cinética de la reacción. También se puede determinar el área de fijación de la proteína al ADN mediante deleciones del promotor. En la figura los constructos *A* y *B* producen una banda de retraso, mientras que, si se elimina la secuencia ATTGG, la banda desaparece (constructos *C* y *D*), lo que demuestra que el factor de transcripción se une a la mencionada secuencia.

La transcripción se inicia tras fijarse los factores de transcripción al ADN e interaccionar con otros factores o con la ARN polimerasa II y activar la enzima. Para investigar qué área del factor de transcripción es responsable de la activación de la transcripción, se han llevado a cabo varios experimentos, en los que se han creado proteínas híbridas que contienen el lugar de fijación al ADN y diferentes áreas de otros factores de transcripción.

Se ha podido demostrar así que existen dominios de las proteínas, generalmente diferentes del que se une al ADN, que son los responsables de la activación. Comparando los dominios de activación de distintos factores de transcripción, se ha comprobado que la mayoría poseen una proporción elevada de aminoácidos acídicos que forman una hélice alfa con las cargas negativas en su superficie. Esto no ocurre en todos los casos; el dominio del factor Sp1, por ejemplo, contiene una región rica en glutamina, mientras que el del CTF/NF1 es rico en prolina. Los dominios de activación son los lugares por donde la proteína se une probablemente a otros factores de transcripción o a la ARN polimerasa II.

A pesar de que la mayoría de los factores de transcripción actúan de una manera positiva, se han descrito casos en los que el factor de transcripción ejerce un efecto inhibidor. A veces se encuentran factores de transcripción con efectos opuestos, unos que inducen la transcripción y otros que la bloquean; esto último ocurre uniéndose a la misma o a diferente región del promotor.

Demos algún ejemplo. El promotor del gen del interferón alfa posee una secuencia AAGTGA, que es necesaria para la transcripción. Utilizando esta región de sonda se ha clonado un factor de transcripción llamado IRF-1 (de "Interferon Regulatory Factor-1"), que se une al ADN e induce la transcripción del gen para el interferón alfa. Recientemente se ha identificado otro factor de transcripción, IRF-2, de estructura similar al IRF-1. (Entre las dos proteínas hay un 62 por ciento de identidad en la región aminoterminal, que es por donde se une al ADN.) A pesar de que el IRF-2 se une a la misma región del ADN, no estimula la transcripción; la inhibe del todo.

Algunos de los factores de transcripción son específicos de tejidos. Es decir, se dan sólo en un tipo celular y están en relación con la expresión específica de genes por dichas células. Por ejemplo, el GHF-1 es un factor de transcripción necesario para la expresión de la hormona del crecimiento y sólo se encuentra en las células de la hipófisis anterior, única sede de producción de dicha hormona.

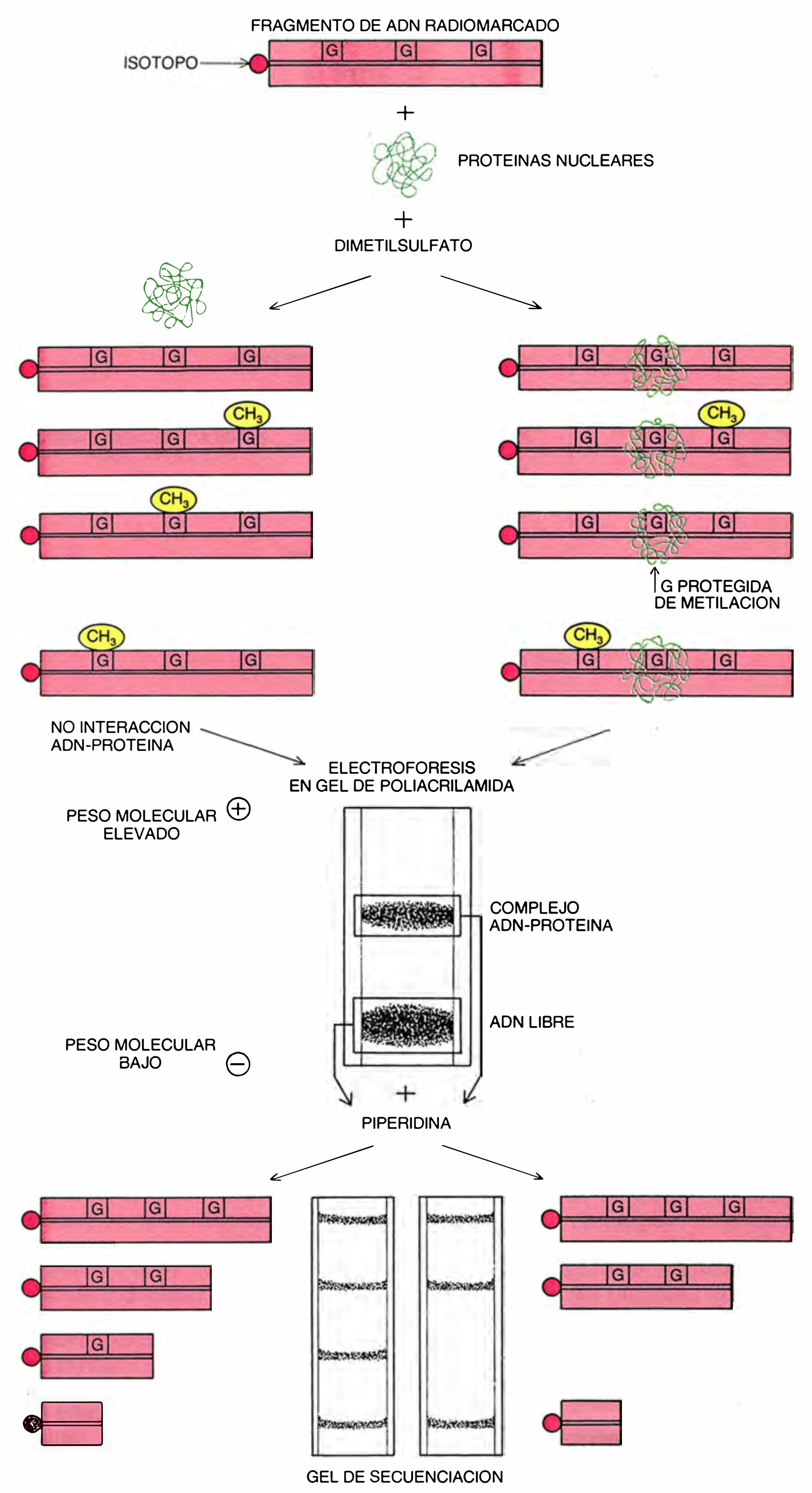
La mayoría de los factores de transcripción se expresan, sin embargo, en todos los tejidos del organismo. En algunos casos en los que se han purificado estas proteínas, se ha observado que gozan de la capacidad de fijarse con la misma afinidad a diferentes regiones de ADN con secuencias de nucleótidos escasamente parecidas. En estos casos, la misma área de la proteína que se fija al ADN tiene varias especificidades diferentes. Tenemos un ejemplo de ello en la proteína OBP 100 ("100 Kilodaltons Octamer Binding Protein"), que se une a la secuencia ATGCAAAT del promotor del gen de las inmunoglobulinas; puede, además, unirse específicamente a la secuencia reguladora TAATATAT del gen IE del virus del herpes, a pesar de contar con sólo tres bases homólogas. En ciertos casos, como el del receptor de los glucocorticoides, según el promotor sobre el que actúen, pueden tener un efecto positivo o negativo en la transcripción, aunque sea la misma área de la proteína la que interaccione con el ADN.

No sólo la misma proteína puede unirse a distintas regiones reguladoras, sino que, además, pueden unirse,

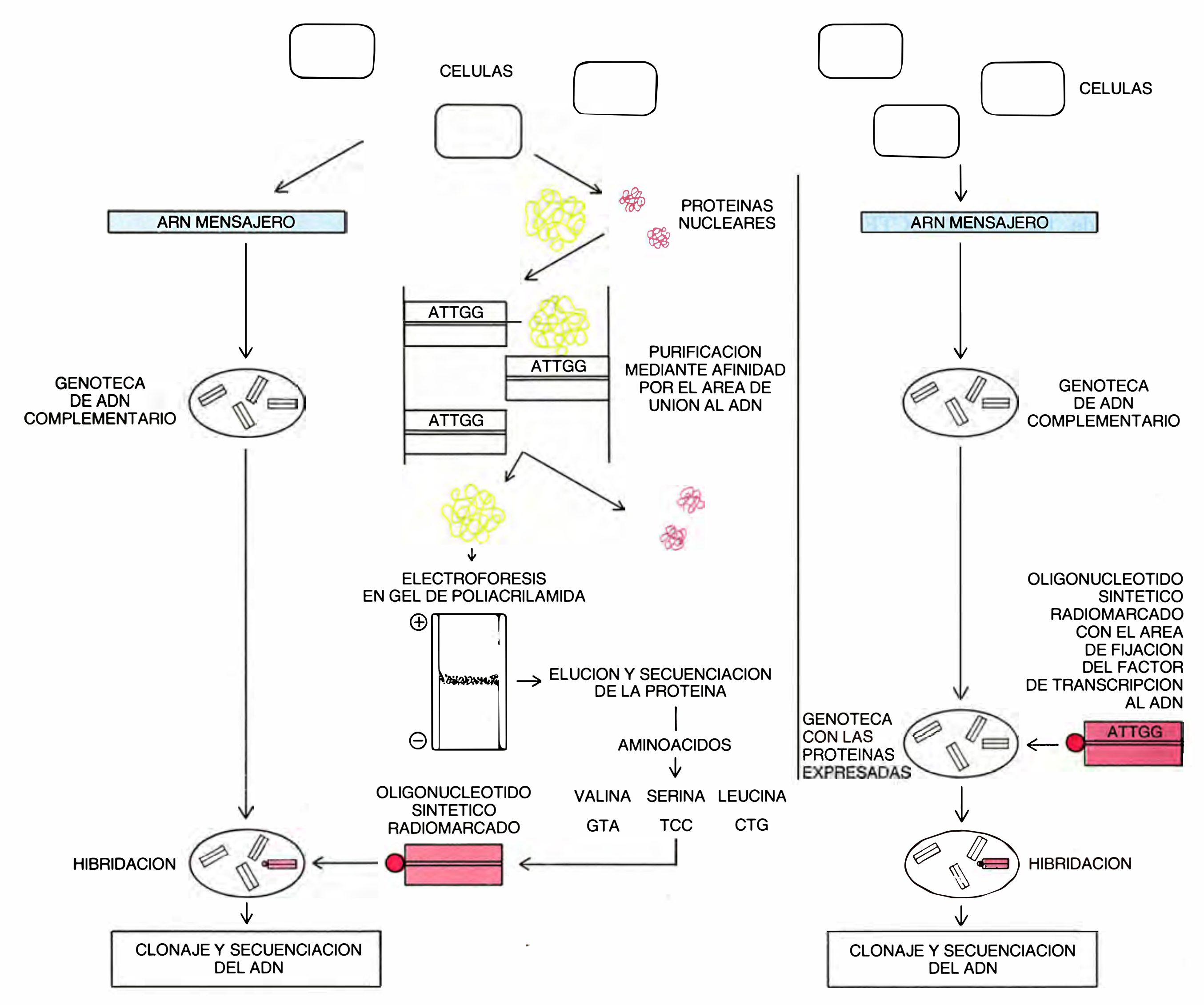
a la misma región reguladora, distintos factores de transcripción. Un ejemplo de esto lo constituye el pentanucleótido CCAAT, que se encuentra en muchos genes a 50 o 100 bases del lugar donde se inicia la transcripción. Se han purificado y clonado los factores de transcripción CTF/NFI y C/EBP; en ambos, la secuencia CCAAT se erige en lugar de fijación al ADN. Estos factores están codificados por dos genes distintos y utilizan motivos diferentes para unirse al ADN. Se ha hallado un tercer grupo de proteínas, de características fisicoquímicas peculiares, que reconoce la secuencia CCAAT. ¿Habrá que admitir la existencia de una familia de factores de transcripción que reconozca como lugar de fijación al ADN la misma secuencia?

En ciertos casos, se requiere la interacción de dos proteínas diferentes para obtener una fijación eficiente al ADN. Un ejemplo es el factor de transcripción que se une a la secuencia Y del promotor de los antígenos de clase II del complejo mayor de histocompatibilidad, proteínas que se expresan en la superficie celular, que, entre otras cosas, sirven para distinguir lo propio de lo extraño. De hecho, la secuencia Y contiene el pentanucleótido CCAAT, al que se unce el factor de transcripción. Mediante cromatografía hemos podido separar dos componentes de ese factor, uno aniónico (Factor A) y otro catiónico (Factor B), que, una vez purificados, tenían un peso molecular de 32 y 42 kilodaltons, respectivamente. Sólo el componente A es capaz de unirse, y débilmente, al ADN; ahora bien, si se añade el componente B, su afinidad (fuerza de unión) se incrementa unas 20 veces.

Se ha descubierto recientemente que algunos factores de transcripción pueden unirse con otros e inhibir su interacción con el ADN, lo que puede explicar algunas observaciones *in vivo*. Durante los procesos inflamatorios en las articulaciones (artritis), los macrófagos liberan interleucina 1 o el factor de necrosis tumoral; estos mediadores actúan sobre los fibroblastos sinoviales que, en respuesta,



5. LOCALIZACION EXACTA de la fijación de la proteína al ADN, determinada mediante el método de "improntas". El ADN radiomarcado en un extremo se trata, en presencia de extractos nucleares, con dimetilsulfato, agente que metila las guaninas (añade un grupo CH_3) de una forma aleatoria; es decir, una sola G por molécula de ADN. Si el factor de transcripción se une a una G, la protegerá de ser metilada por el dimetilsulfato. Tras la migración en geles, el ADN libre, o en forma de complejo con las proteínas, es aislado y tratado con piperidina. Este agente fracciona las moléculas de ADN por allí donde las G están metiladas: corta en un solo lugar cada molécula de ADN creando un número igual de fragmentos de muy diversos tamaños. Al hacer migrar el ADN así tratado por un gel de secuenciación (de un espesor de 1 milímetro o menos), aparecen múltiples bandas que habrán migrado según el tamaño. Si en los extractos añadidos al ADN hay una o varias proteínas que se fijan con suficiente afinidad (intensidad), lo protegerán de la metilación y posterior digestión, con lo cual el ADN no se fraccionará. En este caso, a diferencia del control, no se verán bandas en la posición del gel correspondiente al fragmento de ADN protegido de la digestión; ello señala la posición exacta de contacto entre la proteína y el ADN.



6. CLONACION Y SECUENCIACION de los nucleótidos de ADN del gen que determina la proteína reguladora del ADN. Ofrecemos dos modelos. En el de la izquierda, la primera etapa consiste en purificar la proteína a partir de grandes cantidades de extractos celulares. Esto se lleva a cabo mediante métodos de cromatografía, esto es, aprovechando las propiedades físico-químicas de la proteína, tales como el peso molecular o la carga, que la distinguen del resto de las proteínas de los extractos nucleares. La proteína se suele obtener totalmente pura tras el paso a través de una columna de oligonucleótidos sintéticos (fragmentos de ADN) que contienen la secuencia (ATTGG en el ejemplo de la figura) del ADN a la que se une la proteína, lo que se habrá determinado previamente mediante el método de "improntas". Una vez purificada la proteína, conoceremos la secuencia de aminoácidos que la componen. Se preparan luego oligonucleótidos sintéticos que sirven de sonda para la clonación del ADN. La sonda radiomarcada se hibrida con el ADN de una genoteca (ADN total de la célula). Ello nos dará el ADN que determina la proteína que buscamos.

liberan colagenasa, enzima capaz de destruir el cartílago y causar el malfuncionamiento de la articulación, que se acompaña del dolor e hinchazón característicos de la enfermedad. Una de las drogas utilizadas para combatir la artritis son los glucocorticoides, entre cuyos mecanismos de acción se cuenta el de inhibir la secreción de colagenasa.

El promotor del gen de la colagenasa contiene una región, la AP-1, que es el elemento que induce su expresión. El factor de transcripción que reconoce la región AP-1 es un complejo formado por las proteínas *c*-Jun y *c*-Fos, cuya activación induce la expresión de colagenasa. Los estudios llevados a cabo por distintos laboratorios han demostrado que la proteína *c*-Jun puede unirse al receptor de los glucocorticoides (factor de transcripción localizado en el núcleo y que, inducido por los glucocorticoides, constituye el mediador de su actividad). Esta interacción hace que la proteína *c*-Jun no se una al *c*-Fos, con lo cual el factor de transcripción AP-1 no se fija al promotor de la colagenasa y esta enzima no se secreta en la articulación.

Una vez descritas las propiedades de los factores de transcripción, importa abordar su modo de operación. El papel principal de muchos factores de transcripción consiste en activar un gen determinado en un tejido específico o en respuesta a un estímulo en particular. Deben existir, por tanto, algunos mecanismos que garanticen que un factor dado intervenga sólo en el tejido específico o se active en respuesta exclusiva a cierta señal. En la mayoría de los casos, se fijan sobre el promotor más de una proteína o grupos de proteínas. Estos grupos de factores de transcripción permiten a los promotores responder a distintas señales intracelulares. Por citar un ejemplo, el promotor de la metalotionina posee regiones separadas que, ante los glucocorticoides, metales u otros estímulos, reaccionan induciendo la transcripción del gen.

Uno de los mecanismos de regulación de los factores de transcripción consiste en sintetizar el factor en respuesta a un determinado estímulo: la presencia del factor de transcripción condiciona que tenga lugar la expre-

sión del gen. La proteína Oct-2 es un factor de transcripción necesario para la activación del gen de las inmunoglobulinas; sólo se presenta cuando las células producen inmunoglobulinas.

La regulación de la síntesis del factor de transcripción —que, no se olvide, es una proteína—, se lleva a cabo actuando sobre la transcripción del gen que determina el factor. ¿Cómo se controla el gen que produce el factor de transcripción que controla otro gen? Poco se sabe a este respecto. Lo habitual es que el promotor de los factores de transcripción posea una secuencia sobre la que actúa el propio factor de transcripción, produciéndose una autorregulación. Por ejemplo, el gen que codifica para el factor de transcripción GHF1, que regula la hormona del crecimiento en la hipófisis anterior, contiene en el promotor una secuencia sobre la que se fija el propio GHF1 induciendo la expresión del gen. Además, en el promotor existe un área que se regula por los niveles intracelulares de AMP cíclico (mediador de señales entre el espacio extracelular e intracelular), molécula que inicia, probablemente, la transcripción, en tanto que GHF1 perpetúa la transcripción de su propio gen.

Sin embargo, el mecanismo más común de regulación de los factores de transcripción no depende de su síntesis. Están siempre presentes, aunque no activos. En algunos casos, los factores de transcripción se activan tras sufrir ciertas modificaciones: fosforilación o glicosilación. Entonces sí son capaces de inducir la transcripción; así, el factor de transcripción por choque térmico, o factor HSTF, que induce la expresión de una serie de proteínas cuando las células se incuban a temperaturas superiores a los 37 °C. En estas condiciones, este factor se fosforila e induce la transcripción. Si la proteína se trata con fosfatasa alcalina, que elimina la fosforilación, el factor de transcripción pierde su actividad.

En otros casos, la unión de un agente activador con el factor de transcripción produce modificaciones estructurales de éste haciéndole activo. Citemos, a modo de ejemplo, el factor ACE1 que activa el gen de la metalotionina en respuesta al cobre. La proteína en presencia del cobre experimenta una modificación de su estructura que le permite fijarse al ADN.

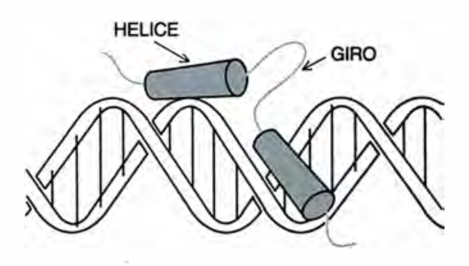
Se ha descrito un mecanismo diferente en la familia de los receptores nucleares para los esteroides y hormonas tiroideas; la presencia de la hormona desplaza a un factor intermediario que bloquea la unión del factor de transcripción al ADN, permitiéndole actuar. En ausencia de hormona, el factor intermediario inhibe la fijación del factor de transcripción al ADN. Es decir, en este caso, el factor de transcripción no es activado por modificaciones estructurales, sino por el fracaso de un proceso: la suspensión de la interacción inhibidora entre proteínas.

El control de la expresión de los genes se lleva también a cabo mediante factores de transcripción que actúan de forma negativa y que se han denominado represores. Estas proteínas operan de distintas maneras. La forma más simple: competir con un factor de transcripción que tenga un efecto positivo por un mismo lugar de fijación al ADN. Hemos aludido antes al promotor del interferón alfa, en el que el represor IRF-2 se une al ADN y bloquea la interacción del IRF-1 con el ADN y, por tanto, la transcripción. Cuando las células se infectan con virus, el represor IRF-2 se inactiva, lo que permite al IRF-1 unirse al ADN y activar la expresión del gen del interferón alfa.

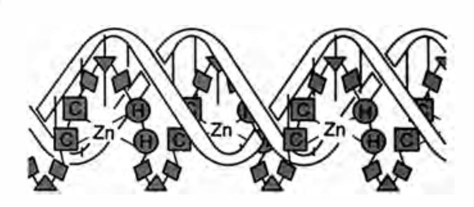
En otro tipo de inactivación, conocido por mecanismo de silenciamiento ("quenching"), las proteínas activadoras y represoras se unen a zonas adyacentes del ADN y, mediante interacciones entre ambas, la proteína represora bloquea la función de la activadora. Es decir, impide su operación sin oponerse a su unión con el ADN. Los represores pueden también ejercer su efecto directamente uniéndose a un lugar apropiado del ADN e inhibiendo la transcripción. Por último, la represión puede constituir el resultado, paradójico, de una expresión excesiva de proteínas activadoras. A este mecanismo se le ha llamado secuestro ("squelching"). Algunos activadores interaccionan con otros factores de transcripción; si estos últimos se producen en cuantía excesiva, impiden que el activador se una al ADN que le conviene, y ello comporta la represión de la expresión de los genes.

Hemos visto que existen distintos mecanismos para activar o inhibir la expresión de los genes. Hemos visto también que muchos de los factores de transcripción pueden actuar en distintos promotores reconociendo secuencias de ADN que no son idénticas; a su vez, la misma secuencia puede ser reconocida por distintas proteínas. ¿Cómo funciona este sistema? Aún no tenemos una respuesta exacta y tendremos que esperar varios años antes de resolver el problema, si bien se especula sobre las posibles interacciones entre los distintos factores de transcripción. Queda, asimismo, por resolver de qué manera las señales recibidas en la superficie celular llegan al núcleo, donde activan los factores de transcripción para iniciar la expresión de los genes.

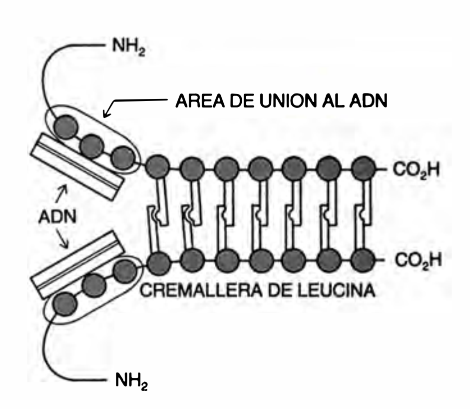
HELICE-GIRO-HELICE



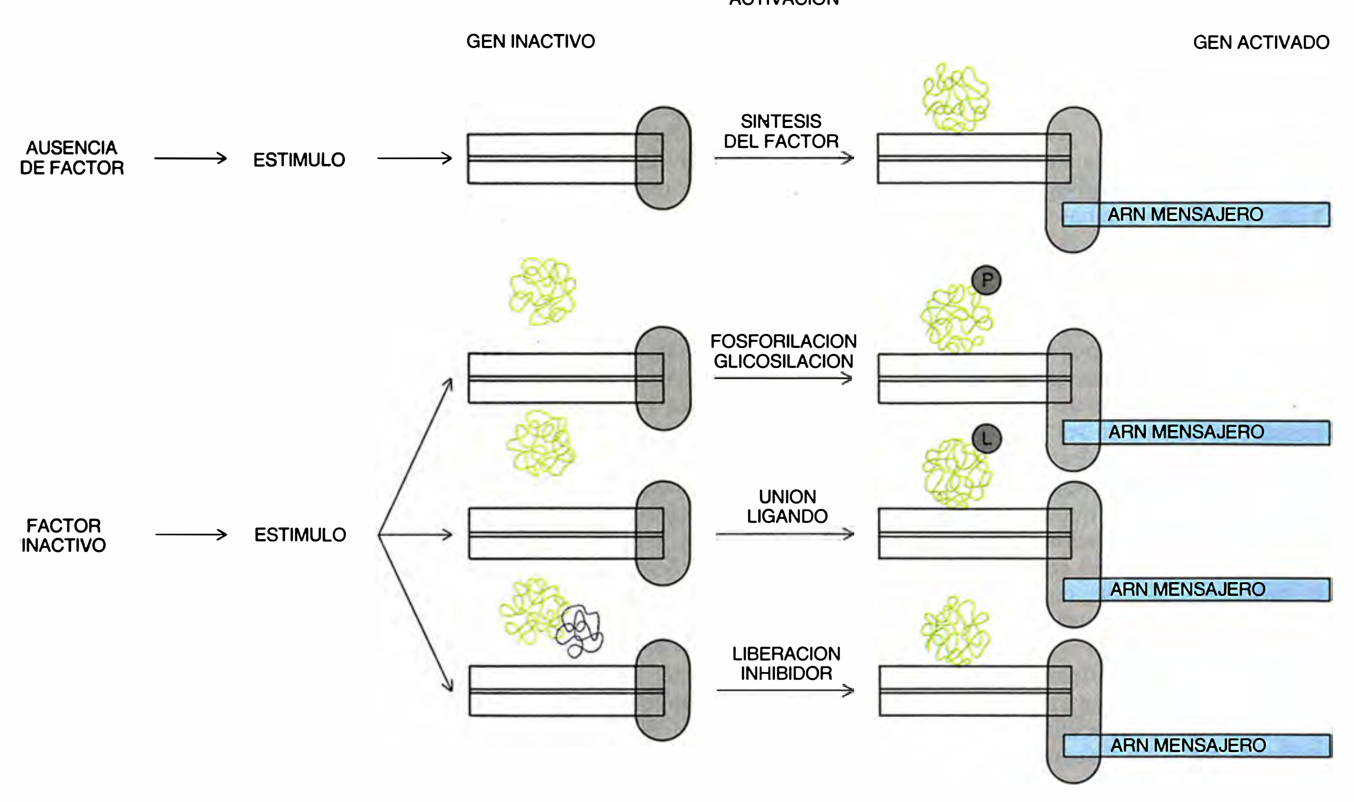
DEDOS DE ZINC



CREMALLERA DE LEUCINA



7. INTERACCION entre factores de transcripción y ADN, puesta de manifiesto en los trabajos de clonación acometidos en los últimos años. La primera estructura es la llamada hélice-giro-hélice (*a*). Los aminoácidos de la "hélice de reconocimiento" tienen un contacto directo con las bases del ADN. La otra hélice se coloca alrededor y se une al ADN de forma aleatoria. El motivo en "dedos de zinc" (*b*) consiste en series repetidas de aminoácidos que contienen cada una dos cisteínas y dos histidinas. Entre estas dos parejas se inserta un átomo de zinc y los aminoácidos entre ellas se proyectan como un dedo cuya punta se halla en contacto directo con el ADN. En la "cremallera de leucina" (*c*) cada siete aminoácidos hay un residuo de leucina; en esta estructura helicoidal, las leucinas se encuentran cada dos giros en el mismo lado de la hélice. Las cremalleras de leucina no se unen directamente al ADN, sino que facilitan la dimerización (unión de proteínas) mediante interdigitalización de las dos hélices que contienen leucina. La dimerización produce una estructura proteica que se une al ADN a través de una región adyacente.



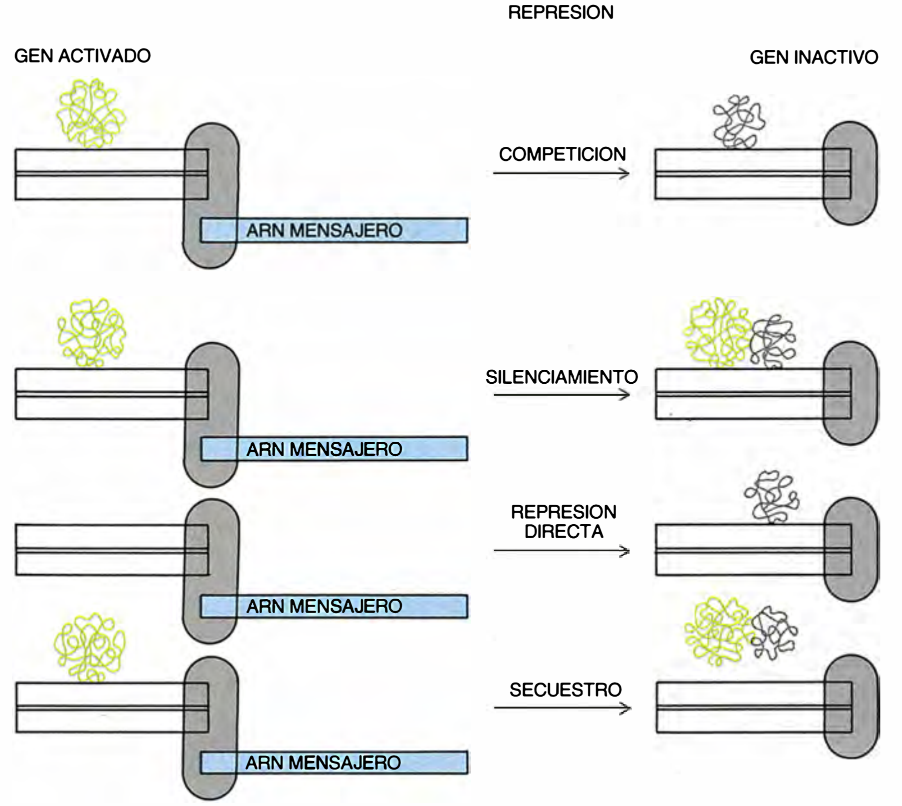
8. LOS FACTORES DE TRANSCRIPCION controlan la expresión de los genes utilizando diversos mecanismos. El primero de ellos es la síntesis del factor tras ser estimulada la célula; la presencia del factor condiciona la expresión del gen. Ahora bien, en la mayoría de los casos el factor de transcripción está presente, por eso a veces se requiere que el factor se active tras modificaciones estructurales (fosforilación o glicosilación) o tras unirse a un ligando. En otros casos, es preciso que se libere un inhibidor del factor, pudiéndose entonces unir al ADN. Los factores de transcripción uti-

La interacción entre los factores de transcripción con el promotor tiene por finalidad comenzar la transcripción. Según lo señalado, el promotor se encuentra a corta distancia del origen de la transcripción donde opera la enzima polimerasa II. La regulación se lleva a cabo mediante multiplicadores ("amplificadores") que se encuentran, muchas veces, bastante lejos del promotor. Una de las explicaciones de esta observación es que el ADN puede doblarse de suerte que las proteínas que se unen a las áreas de amplificación entren en contacto directo con las proteínas que se unen al promotor. Tenemos un ejemplo que ilustra estas interacciones en el promotor de los antígenos de histocompatibilidad de clase II. Los genes que determinan estos antígenos poseen en el promotor dos secuencias que caen a unas 100 bases del lugar donde se inicia la transcripción; entre tales secuencias —X e Y se han llamado— median 20 bases, por cuya razón ambas se encuentran en el mismo lugar del giro del ADN. Curiosamente, estas dos secuencias, aunque en posición invertida, se hallan también a una distancia de 1500 bases de las primeras (Y' y X'). Las proteínas que se unen a las secuencias X o Y son las mismas que se unen a las secuencias X' e Y'. Esto sugiere que el ADN, al doblarse sobre sí mismo, puede hacer que los factores de transcripción unidos a X interaccionen con los unidos a X' y los unidos a Y con los de Y' (dimerización).

¿Cuáles son las aplicaciones prácticas obtenidas del estudio de los factores de transcripción? Lógicamente, el conocer el mecanismo de transcripción de los genes nos permitiría modular su expresión, de cuyo máximo interés el lector se hará en seguida una idea a través de varios ejemplos.

Existe un tipo de inmunodeficiencia (o de enfermedad en que los pacientes, privados de defensas inmunes, contraen por ello toda suerte de infecciones) en el que las células no expresan los antígenos de histocompatibilidad de clase II. En estos sujetos se ha demostrado que carecen de uno de los factores de transcripción que se unen al promotor, concretamente al área X. En otros casos, la interacción con el promotor es defectuosa a pesar de la presencia de los factores de transcripción. La hemofilia B constituye una enfermedad hereditaria en que falta el factor IX del sistema de coagulación, lo que origina problemas hemorrágicos. Se ha visto que, en algunos de estos pacientes, el promotor del gen que determina el factor IX presenta una sustitución de una adenina por una guanina en la base -13. El factor de transcripción C/EBP se fija en esta región, pero no se une a la secuencia mutada; al no hacerlo, no se induce la producción de factor IX y consiguientemente se origina la enfermedad.

Uno de los campos donde el estudio de los factores de transcripción cobra mayor interés es el de la investigación oncológica. La proliferación y diferenciación celular está normalmente controlada por factores de transcripción que actúan estimulando o inhibiendo, y modulan así las necesidades del organismo. En el cáncer se produce un desajuste de ese equilibrio; resulta entonces una proliferación celular anárquica, por un exceso de factores que estimulan la proliferación o por defecto de inhibidores. La noción de oncogén se interpretó inicialmente de una forma errónea. Hace unos años se vio que las células infectadas por ciertos virus se transformaban en células tumorales. Esto hizo pensar que los virus introducían en las células nuevos genes, por lo que se les llamó oncogenes, que eran



lizan diversos mecanismos para inhibir la expresión de los genes. Uno consiste en competir por el lugar de fijación al ADN de los factores activadores. En virtud del otro mecanismo, el de silenciamiento, se une al ADN en una región adyacente e inhibe la función del activador. Los represores pueden unirse al ADN y bloquear la actividad directamente.

los responsables de la producción de tumores. Se descubrió luego que estos oncogenes se daban normalmente en las células y contribuían, muchas veces, al proceso de división celular. Curiosamente, muchos de estos oncogenes son factores de transcripción que actúan induciendo la expresión de genes necesarios para la multiplicación celular. Si los oncogenes se activan tras la infección por los virus, la proliferación celular se descontrola y se desencadena el proceso tumoral.

No se trata sólo de una idea comprobada en el laboratorio; se ha descrito también en ciertas situaciones de la patología humana. Por ejemplo, en el 96 por ciento de los casos de leucemia mieloide crónica, se presenta una alteración cromosómica conocida por "cromosoma de Filadelfia". En este caso, el gen *c-abl* abandona su ubicación normal en el cromosoma 9 para instalarse en el 22, inserto en una región llamada *bcr*. Esta yuxtaposición (*bcr/c-abl*) da origen a un factor de transcripción que hace proliferar incontroladamente las células de la médula ósea y provoca la leucemia. El gen originado por la fusión *bcr/c-abl* introducido en modelos animales ocasiona una enfermedad en el ratón equivalente a la leucemia mieloide crónica humana. Cierto tipo de leucemias agudas linfoblásticas (pre-B) se caracteriza por una alteración cromosómica en la que parte del cromosoma 1 pasa al 19 (translocación) y se inserta en una región de este último, donde se localiza el gen E2A que codifica para un factor de transcripción que se une al promotor de las inmunoglobulinas. En virtud de esa translocación, se forma un nuevo factor de transcripción que estimula la producción de linfocitos B, con resultado de leucemia. El retinoblastoma es un tumor que, en los niños que lo padecen, causa la destrucción de la retina. En este caso se ha encontrado una mutación en el cromosoma 13 en una región donde se aloja el gen del retinoloblastoma. Las células de estos pacientes no sintetizan el producto de este gen, la proteína *p*105-RB; es ésta un factor de transcripción que actúa como represor de la proliferación celular. Al faltar dicho factor, las células proliferan sin orden y se produce el tumor.

Hemos visto algunos ejemplos de la implicación directa de factores de transcripción en la aparición de algunas enfermedades. En otros casos, hay pruebas que hablan en favor de un papel importante de estos factores en el desarrollo de la enfermedad, aunque se desconoce su mecanismo íntimo de acción. Nos referimos a ciertas enfermedades autoinmunes, así se llaman aquellas en que se produce una reacción del sistema inmune contra algunas células o tejidos del propio organismo. Se ha demostrado que muchos de los mal llamados anticuerpos contra el ADN que se detectan en la sangre o en los tejidos de los pacientes que sufren lupus eritematoso diseminado, síndrome de Sjogren, artritis reumatoide, dermatomiositis, esclerodermia, entre otras, no están realmente dirigidos contra ADN, sino contra diversos factores de transcripción. En el momento actual se desconoce de qué forma estos anticuerpos modifican la actividad funcional de los factores de transcripción y provocan enfermedades. Por último, en un terreno más teórico que experimental, pues se carece de pruebas, se ve en el envejecimiento la imagen invertida de los procesos tumorales. En este caso, la intervención de los factores de transcripción activadores caería, en tanto que se reforzaría la de los inhibidores. Los únicos datos de que disponemos en el momento actual son la relación que existe entre el envejecimiento y los procesos tumorales y algunos estudios provisionales con modelos experimentales. Pero es el comienzo. El futuro nos habrá de deparar muchas sorpresas sobre los factores de transcripción y su modo de controlar la expresión de los genes.

BIBLIOGRAFIA COMPLEMENTARIA

EUKARYOTIC TRANSCRIPTIONAL REGULATORY PROTEINS. P. F. Johnson y S. L. McKnight en *Annual Review of Biochemistry*, vol. 58, págs. 799-839; 1989.

RNA POLYMERASE B (II) AND GENERAL TRANSCRIPTION FACTORS. M. Sawadogo y A. Sentenac en *Annual Review of Biochemistry*, vol. 59, págs. 711-754; 1990.

DNA RECOGNITION BY PROTEINS WITH THE HELIX-TURN-HELIX MOTIF. S. C. Harrison y A. K. Aggarwal en *Annual Review of Biochemistry*, vol. 59, págs. 933-969; 1990.

ACTIVATORS AND TARGETS. M. Ptashne y A. F. Gann en *Nature*, vol. 346, págs. 329-331; 1990.

TOO MANY TRANSCRIPTION FACTORS: POSITIVE AND NEGATIVE INTERACTIONS. M. Karin en *The New Biologist*, vol. 2, págs. 126-131; 1990.

MULTICOMPONENT TRANSCRIPTION FACTOR COMPLEXES: THE EXCEPTION OR THE RULE. P. E. Shaw en *The New Biologist*, vol. 2, págs. 111-118; 1990.



Diseño y construcción de los arcos primitivos

El arco primitivo, que fue durante milenios el útil más resolutivo de cazadores y guerreros, sufrió múltiples modificaciones que reflejaban las diversas necesidades funcionales del arquero

Edward McEwen, Robert L. Miller y Christopher A. Bergman

Pocos serán quienes discutan que la invención de la rueda y el dominio del fuego se cuentan entre los acontecimientos históricos a destacar. Sin embargo, la creación del arco se deja pasar, más de una vez, inadvertida. Desde el Paleolítico hasta la aparición de las armas de fuego en el siglo XVI, el arco fue, amén de útil venatorio de primer orden, la principal entre las armas de guerra. Esencial se mostró para las tribus nómadas de Asia Central que conquistaron tierras y fundaron dinastías en China, así como para los sitiadores de castillos en la Europa medieval. Este arma, que encontramos en la inmensa mayoría de las culturas, muestra en su constitución variaciones importantes y de índole práctica, que van desde toscas ramas con una cuerda tirante hasta ingenios mecánicos de evidente refinamiento.

En sus rasgos esenciales, el arco es un resorte de dos brazos sujeto por los extremos mediante una cuerda que lo mantiene bajo tensión. Al tensar un arco, el lomo de éste, que es la curva exterior, queda sometido a un esfuerzo de tracción y el vientre, que es la curva interior, a fuerzas de compresión. Para que no se rompa y se impulse la flecha adecuadamente, el arco debe adaptarse a esas fuerzas. Una vez tenso del todo, en sus brazos se almacena energía potencial. Al soltar la cuerda, esa energía se transfiere a la flecha, que sale proyectada al aire.

Inmortalizado por la leyenda y la historia, el arco inglés es quizás el más conocido. No obstante, pese a su capacidad letal, el arco inglés no resultaba, por su tamaño, de manejo práctico para disparar, por ejemplo, desde un caballo. Por excavaciones arqueológicas y nuestros propios experimentos con réplicas, sabemos que los constructores de arcos primitivos adaptaban el arma a sus necesidades particulares; así nacieron innumerables variantes ingeniosamente diferentes. Los teton lakota (o sioux) acortaron los arcos para facilitar su manejo a caballo; los hunos combinaron materiales diferentes para hacer arcos pequeños, aunque extraordinariamente vigorosos, cuyas flechas traspasaban la armadura enemiga.

1. EN ESTE GRABADO sobre chapa de cobre, que data de 1765, arqueros manchúes, armados de arcos compuestos, combaten a guerreros tártaros. Encargada por el Emperador Chien-Lung, la pieza conmemora la conquista de Tartaria a mediados del siglo XVIII.

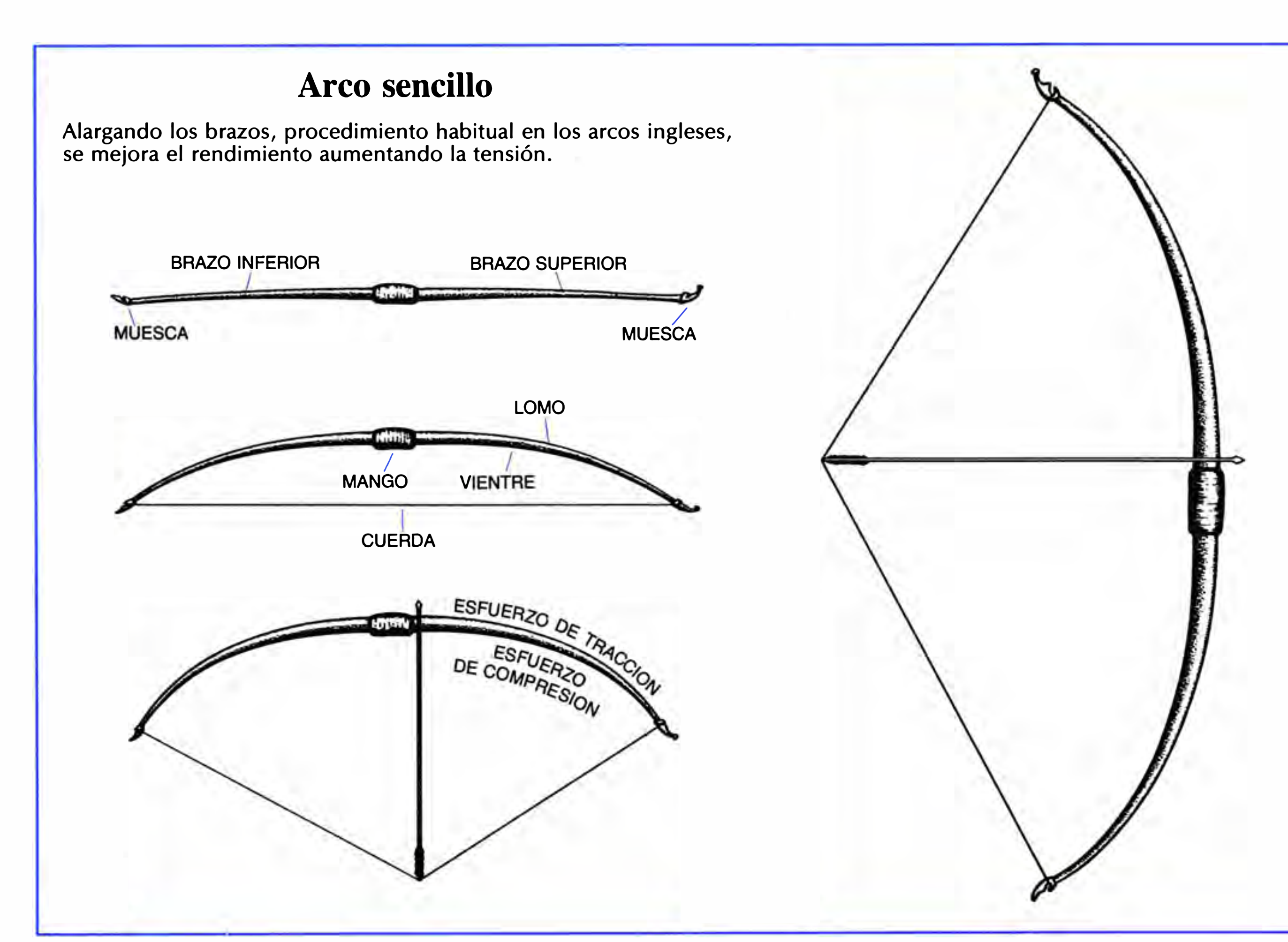
A partir del Paleolítico el diseño de los arcos siguió, en líneas generales, tendencias diferentes en Europa y Asia. Ninguna de ellas puede considerarse intrínsecamente mejor que la otra. Cada diseño representa una solución posible al problema de lanzar, con tino, un proyectil pequeño y liviano, dotado de poder de penetración.

Los distintos tipos de arco no aparecieron de golpe. Su diseño parece haber seguido un proceso gradual de modificaciones, que abarca muchos milenios y culturas prehistóricas. Aunque hubo expertos que creían que el arco inglés fue invento de los anglosajones, normandos o galeses, los investigadores han descubierto antecedentes que se remontan 8000 años atrás por lo menos. Ciertas pruebas respaldan la posibilidad de que los primeros útiles de arquería aparecieran en el Paleolítico Superior temprano (desde el 35.000 hasta el 8000 a.C.). En este artículo rastreamos la evolución del arco desde sus comienzos prehistóricos hasta las modificaciones introducidas, primordialmente en Europa y Asia, hace tan sólo 400 años.

Los testimonios más antiguos sobre el origen del arco podrían ser las púas arrojadizas recuperadas de las sociedades paleolíticas del Viejo Mundo, de las culturas perigordiense y solutrense, que se desarrollaron en lo que hoy es Francia. Delgadas y estrechas por la base, esas púas, datables entre el 28.000 y el 17.000 a.C., encajarían fácilmente en una ranura existente en el astil de una flecha; pero con idéntica probabilidad podrían ser también puntas de dardos arrojadizos.

Los arqueólogos han reunido más pruebas fehacientes en Stellmoor, cerca de la ciudad alemana de Hamburgo: un conjunto de astiles y capiteles de flecha procedentes de una cultura glaciar tardía que existió a principios del noveno milenio antes de nuestra era. No cabe duda de que esos astiles rotos debían emplearse con los arcos. A diferencia de los dardos, que estaban dotados de una cavidad estrecha, o copa, que encaja en el gancho de un mecanismo llamado tiradardos, los astiles en cuestión po-

EDWARD McEWEN, ROBERT L. MILLER y CHRISTOPHER A. BERGMAN han sometido a ensayo arcos procedentes de todas las culturas del mundo. McEwen es el director editorial del *Journal of the Society of Archer-Antiquaries.* Ha traducido y comentado textos medievales persas sobre arquería en la Escuela de Estudios Orientales y Africanos de Londres. Miller trabaja en el Servicio Arqueológico Clover con sede en Northrop. En su labor investigadora aplica el método científico al estudio de la tecnología primitiva, ecología humana y paleoepidemiología. Bergman es arqueólogo principal en la 3D/Enviromental Services, Inc., en Cincinnati, y, durante los últimos diez años, ha recreado la arquería de los aborígenes norteamericanos. Miller y Bergman son doctores por el Instituto de Arqueología de Londres. Los autores desean agradecer a Charles E. Grayson, de Clatskanie (Oregon), y a Frank J. McAvinchey, de 3D/Environmental Services, su ayuda para la redacción de este artículo.



seen ranuras estrechas y rectangulares donde sólo hallarían acomodo cuerdas de arco.

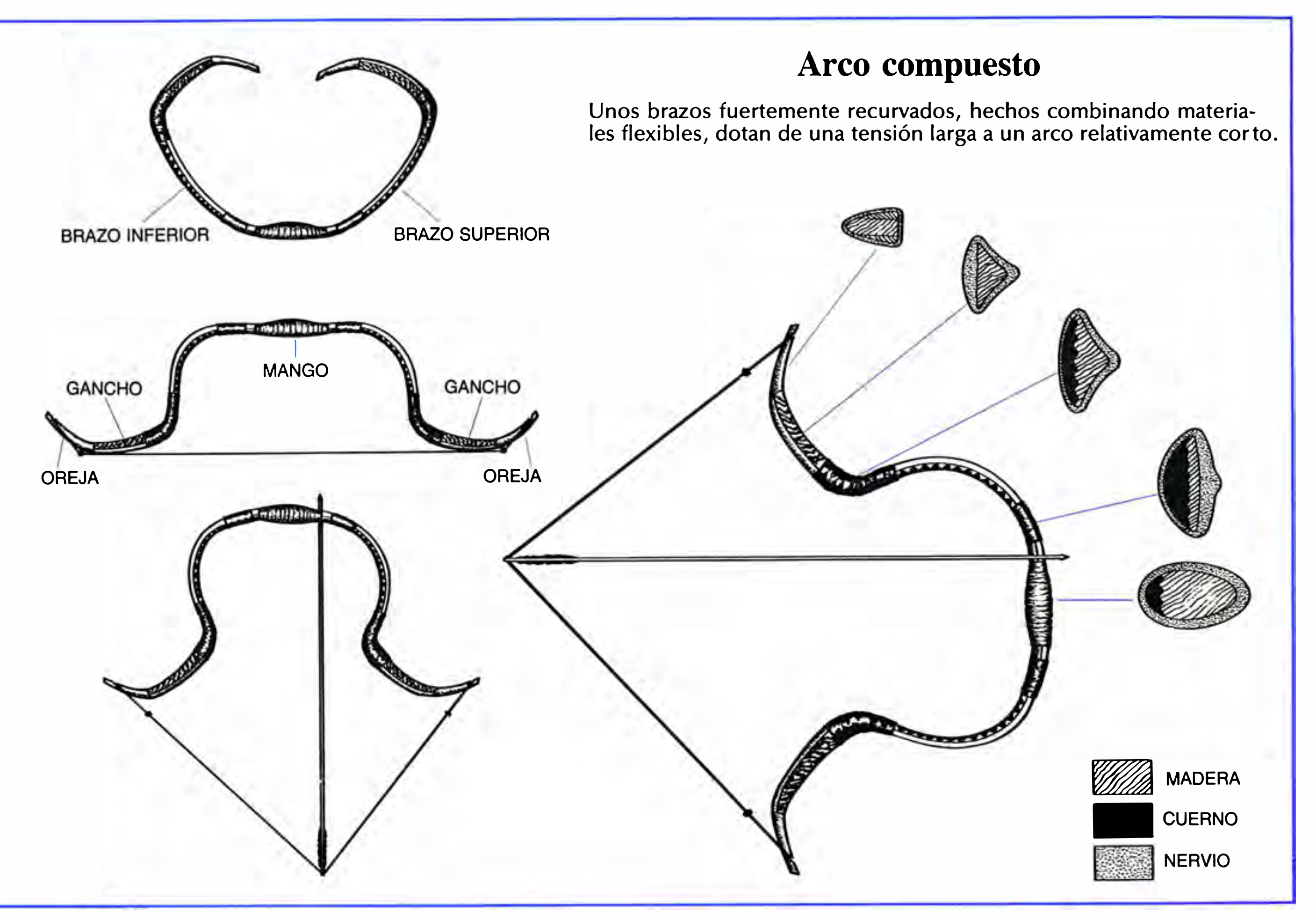
Los arcos más antiguos y enteros recuperados por los investigadores se remontan al 6000 a.C. Conservados en regiones anegadas de Escandinavia, se trata de ejemplares muy elementales, de una sola pieza, principalmente de madera de tejo u olmo. En atención a su materia prima, de un solo tipo, se les denomina arcos sencillos, grupo éste que abarca más miembros.

En el período Mesolítico (del 8000 al 3200 a.C.), apareció en el norte de Europa un tipo de arco más perfeccionado. Los testimonios recuperados del pantano de Holmgaard, en la isla danesa de Zealand, están hechos de varas de olmo. El mango, rígido, mantiene sujetos por compresión unos brazos anchos y aplanados que se van estrechando gradualmente hacia las puntas. Los constructores de arcos primitivos rasquetearon y afinaron esmeradamente el vientre (labrado del arco); lo prueba la curva suave y regular que se forma cuando se arman (es decir, se encordan). Todas estas características contribuyen a que los esfuerzos se repartan de manera uniforme a lo largo del arco, reduciendo con ello la posibilidad de que se rompa y mejorando sus prestaciones.

Los diseñadores de los arcos de Holmgaard los construían, además, de buena longitud, de 150 a 180 centímetros, el valor en el que se movería la arquería medieval. Cuanto más largo es un arco, mayor es la tensión que puede dársele, lo que influye de modo determinante en el rendimiento y capacidad del arma. (En arquería, por tensión se entiende la longitud a la cual puede tensarse un arco; el rendimiento designa la velocidad a la que la flecha se separa del arco, y llámase capacidad de un arco el alcance de aquélla.) Los arcos sencillos cortos, como los que empleaban los lakota y comanches en sus cabalgadas por las llanuras de América del Norte, tienen unas tensiones más cortas, a menudo de sólo 55 a 60 centímetros.

Las etapas que desembocaron en la aparición de los arcos de Holmgaard se desarrollarían, a buen seguro, a través de experimentos sencillos y del conocimiento de las limitaciones que imponían los materiales y los útiles que se empleaban en su manufactura. Acaso sean los arcos sencillos neolíticos los que mejor patentizan el papel que en el diseño de arcos desempeñan materiales y útiles de trabajo. Los autores de este artículo, empleando el mismo juego de herramientas de que disponían los constructores de arcos neolíticos, descubrimos que tales útiles son perfectamente capaces de construir arcos sencillos. Aunque, por supuesto, tales herramientas limitan la perfección y el grado hasta el que puede labrarse la madera. Aduzcamos el ejemplo del arco neolítico inglés de tejo de Meare Heath (Somerset), fechado con radiocarbono en un entorno de 120 años respecto al 2960 a.C.; difiere sustancialmente del conocido arco medieval inglés (construido con útiles de metal) en las formas del lomo y el vientre. Aunque ambos se asemejen en tamaño (unos dos metros), el de Meare Heath presenta un lomo más redondeado y convexo y un vientre más plano que su contrapartida medieval.

El diseño indica que los constructores del arco de Meare Heath aprovecharon la forma de la rama o fusta del árbol original para reducir el tiempo de manufactura. Creemos así que nuestro constructor prehistórico empezó por elegir una rama adecuada o un serpollo maduro del ancho y largo convenientes antes de cortarlo. El



Arco compuesto

Unos brazos fuertemente recurvados, hechos combinando materiales flexibles, dotan de una tensión larga a un arco relativamente corto.

trabajo de carpintería consistiría esencialmente en ahusar el ancho y grosor de la vara. Lo modesto de este trabajo de carpintería es especialmente evidente en el lomo, formado sobre todo por la curva natural externa de la vara original. Un lomo redondeado altera lo menos posible las capas leñosas que hay bajo la corteza, disminuyendo la posibilidad de que el constructor debilite sin querer la textura de la madera cortando las vetas de ésta. La mayoría de los arcos se parten porque el esfuerzo de tracción provoca la rotura del lomo. Muchas veces el punto débil se encuentra allá donde no se ha seguido con cuidado el veteado; entonces, las fibras leñosas se separan y el arco se fisura.

Existe cierta confusión entre los expertos a propósito de los materiales que componen los arcos de tejo primitivos. El tejo se compone de dos capas de madera disímiles, distinguibles a simple vista: una es la albura, exterior y fisiológicamente activa, y la otra es el duramen, central, rojizo y muerto. La albura es elástica y resiste bien la tracción; el duramen es más apto para soportar esfuerzos de compresión. Gad Rausing, de la Universidad de Lund, observó una evidente ausencia de albura en los lomos de los arcos neolíticos de tejo procedentes de yacimientos lacustres suizos, datados entre el cuarto y el tercer milenio a. C. Los investigadores han informado, asimismo, de la ausencia de albura en el arco de tejo de Meare Heath.

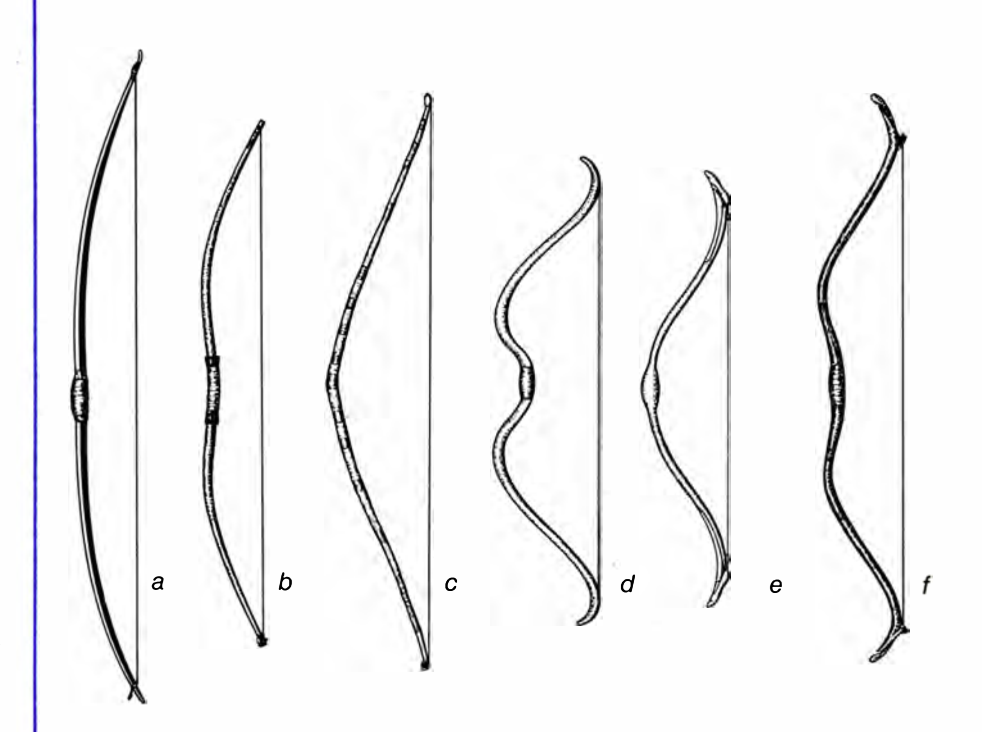
Nuestra experiencia en la construcción de arcos de tejo revela que es muy improbable que esas armas neolíticas se hicieran sólo de duramen. El del tejo, demasiado frágil, no resistiría los grandes esfuerzos de tracción asociados al armado y tensado de un arco. Un arma hecha únicamente de duramen sería muy poco fiable y proclive a romperse en cualquier momento. Quizá los ejemplares del Neolítico se construyeron de madera verde, que sería más elástica, pero el funcionamiento de estos arcos habría sido, como mínimo, deficiente.

El desarrollo de los útiles de metal tras el 2000 a.C. permitió a los artesanos introducir modificaciones de otra índole. Con sus numerosos ejemplares en buen estado de conservación, el arco medieval muestra quizá del mejor modo los tipos de modificaciones que produjeron las herramientas metálicas. Los constructores ingleses manufacturaron sus arcos con varas cortadas de árboles más maduros y mayores que los empleados por sus colegas neolíticos. Como la curvatura externa de un árbol crecido es menor que la de una rama, los constructores tardíos pudieron hacer unos arcos de lomos más planos. Por ejemplo, los arcos que se recuperaron del buque de guerra *Mary Rose*, de Enrique VIII, hundido el 19 de julio de 1545, presentan una sección redondeada, con un lomo levemente aplanado, hecho de albura.

En la época victoriana, las modificaciones tendieron a centrarse en el espesor más que en la anchura de las varas. Esta insistencia es responsable de la sección acusadamente "trapezoidal", tan común en estos arcos. Se consigue mayor rendimiento y capacidad para una menor potencia. (En arquería, por potencia de un arco se entiende la fuerza necesaria para tensarlo.) No obstante, la irregular repartición de las fuerzas a lo largo del estrecho eje longitudinal de un vientre grueso y redondeado predispone el arma a la rotura.

Esos cambios no siempre reflejan una evolución progresiva. En los años treinta y cuarenta, Clarence Hickman, de los laboratorios Bell, Forrest

Nagler, del Colegio de Ingenieros Industriales de Norteamérica, y Paul Klopsteg, del Departamento de Material de Guerra del Ejército Estadounidense, desarrollaron matemáticamente y ensayaron modelos de arcos "ideales". Aquellos modelos están bastante más acordes con los arcos planos de brazos anchos, usados en el Mesolítico y el Neolítico de Europa, que con los arcos tardíos aparecidos en Inglaterra. Nadie sabe por qué los ingleses se decidieron por los brazos estrechos; quizá fuera para economizar la materia prima disponible.



Arcos representativos

Los tipos fundamentales de arco están representados por el arco medieval inglés de tejo (*a*), el Teton-Lakota reforzado con nervios animales (*b*) y cuatro especies de arco compuesto: el angular asiático occidental (*c*), el escita (*d*), el turco del siglo XVII (*e*) y el tártaro de Crimea del siglo XVII (*f*).

Fuera de Europa, el diseño de arcos emprendió un camino diferente. Aunque no hay dudas de que el arco sencillo apareció de manera espontánea e independiente en muchas culturas, las modificaciones más complejas se produjeron en Asia. A diferencia de los europeos, los asiáticos no se concentraron tanto en la arquitectura de los brazos cuanto en los materiales. En particular, los constructores de arcos asiáticos empleaban adhesivos, que extraían de pieles de animales y aletas de peces, para encolar tendones ("nervios animales") a los lomos de los arcos.

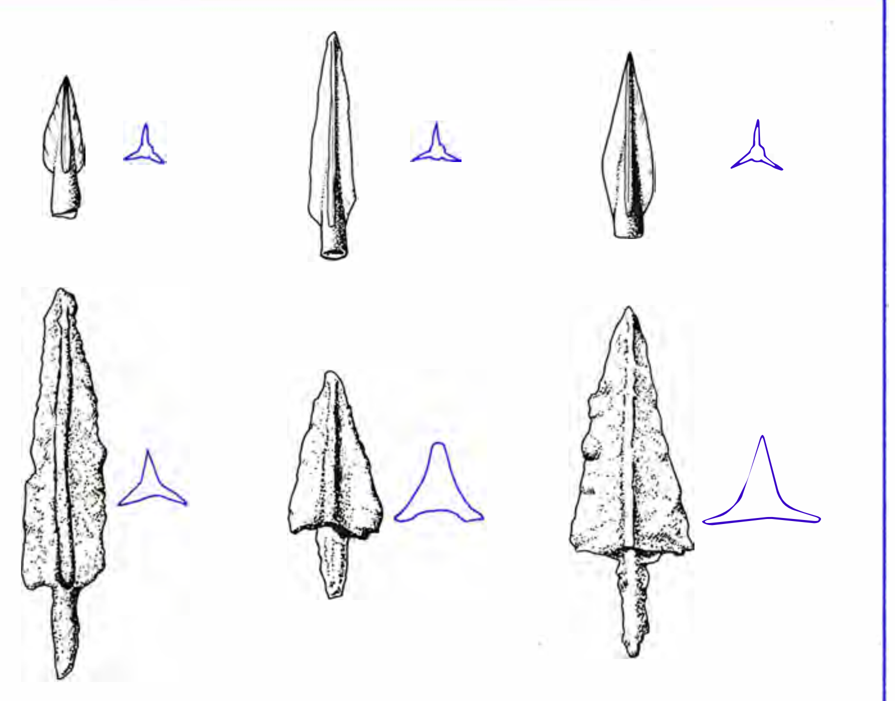
Los tendones poseen una resistencia a la tracción elevada, estimada en torno a los 20 kilogramos por centímetro cuadrado, unas cuatro veces la de los arcos de madera. Esta resistencia permite acortar los arcos de modo importante sin sacrificar la tensión ni aumentar el riesgo de rotura. Muy manejables a caballo, los arcos de brazos cortos reforzados con tendones se emplearon en Asia Septentrional y Extremo Oriente. Algunas tribus indias de las llanuras occidentales de América del Norte desarrollaron y emplearon también arcos de este tipo. (El caballo no es siempre una hipótesis para suponer el desarrollo de arcos reforzados con tendones; los nativos de California los empleaban para cazar a pie en los bosques.)

Una ventaja de primer orden de los arcos reforzados con tendones, o "nervios animales", es que son siempre recurvados: es decir, sus brazos, cuando no están armados, se hallan invertidos. Esto hace que los brazos sufran una fuerza de tracción mayor cuando se arma el arco; así se almacena más energía que en los arcos sencillos. Además, a brazos más cortos mayor eficacia en la transferencia de energía. En efecto, los largos y pesados brazos de los arcos sencillos consumen una buena cantidad de energía cuando se desplazan hacia adelante al soltar la cuerda, resultando con ello que se cede menos energía a la flecha y se derrocha más.

Los constructores de arcos primitivos de Asia Oriental y Occidental no se limitaron a añadir tendones a sus armas. Algunos hubieron de caer en la cuenta de que en la naturaleza existían materiales más vigorosos que la madera. Crearon así un arco más perfeccionado, cuya construcción requería un nivel de destreza importante. Llamado arco compuesto, se trata de un auténtico prodigio mecánico en el que se combinan varios materiales. En su forma clásica, consta de un fino núcleo de madera, con tendones encolados al lomo, y cuerno, generalmente de carabao, con tendones encolados al vientre. Los arqueros modernos dan otros nombres a los arcos de este tipo, entre ellos arcos contraplacados, arcos reforzados, arcos con sobrelomo y arcos mixtos. Aquí el adjetivo "compuesto" lo usamos para referirnos al arco enteramente realizado en cuerno, madera y tendones.

En el arco compuesto se aprovechan bien los materiales que entran en su construcción. Los tendones del lomo soportan los esfuerzos de tracción. El cuerno, con una resistencia máxima de unos 13 kilogramos por centímetro cuadrado (aproximadamente el doble que las maderas duras), soporta las fuerzas compresivas. El cuerno posee, además, un elevado coeficiente de restitución, o aptitud para recuperar su forma original después de haber sido deformado. La flexibilidad de estos materiales permite dotar a los arcos de unos brazos cortos, ligeros y recurvados, capaces de almacenar una gran cantidad de energía cuando están bajo tensión. Además, la gran flexibilidad de los brazos de los arcos compuestos permite tensarlos mucho más con relación a las dimensiones globales del arma. Esta combinación de mayor tensión con brazos más cortos permite descargar las flechas con más velocidad y alcance que un arco sencillo de la misma potencia. Nosotros hemos realizado ensayos que muestran que una réplica de arco compuesto de 27 kilogramos de potencia dispara la misma flecha con el mismo rendimiento que una réplica de arco medieval inglés de tejo de 36 kilogramos de potencia (unos 50 metros por segundo).

Tan sólo la ballesta, inventada hacia el año 500 a.C., puede impulsar un proyectil con más velocidad y alcance. Pero ésta no brilla por su rendimiento; para lograr las enormes potencias que necesita, que pueden rozar la tonelada, hay que equiparla de piezas mecánicas. Así pues, no po-



Puntas de flechas primitivas

Los utensilios de arquería ayudan a los arqueólogos a deducir las necesidades de quienes los utilizaron. Por ejemplo, las puntas de flecha de la fila superior son escitas del siglo III a. C. y están hechas de bronce con una longitud de 25 a 50 mm. Sin embargo, la aparición de la armadura impuso la necesidad de unas puntas más pesadas y largas, hechas de hierro, que pudieran traspasarla. En la fila inferior se muestran ejemplares hunos. A la derecha de cada una se representa su contorno visto de frente.

demos tomarla como un elemento de comparación razonable para un arco ordinario de tiro manual.

Otra ventaja del arco compuesto es que puede mantenerse armado durante períodos prolongados sin efectos adversos. Los arcos sencillos de madera y los reforzados con tendones acostumbran guardarse desarmados entre usos para evitar que "la cuerda siga al proyectil" y trabaje peor.

Desconocemos dónde exactamente se inventó por primera vez, y por quiénes, el arco compuesto. Los registros arqueológicos e históricos sugieren que lo desarrollaron independientemente varias culturas a lo largo del tercer milenio a.C. Más concretamente, las pruebas modernas señalan que el arco compuesto apareció, a la vez y sin relación de dependencia, en Mesopotamia y Anatolia y en las estepas de Asia septentrional.

Augustus Henry Pitt-Rivers, general y arqueólogo inglés, introdujo el término "arco compuesto" a fines del siglo XIX. Consideraba él que estos arcos fueron producto de pueblos que vivían en zonas donde no había buenas maderas para la manufactura de arcos. A primera vista, la argumentación de Pitt-Rivers parece sensata. La verdad es, sin embargo, que el arco compuesto aparece por primera vez en los registros arqueológicos en zonas donde existían amplias reservas de maderas idóneas para arcos. Los egipcios, por ejemplo, construían y empleaban arcos compuestos, pero también producían arcos sencillos de maderas duras: acacia y algarrobo.

Si la carencia de maderas idóneas no fue la causa de la aparición del arco compuesto, hay que suponer que éste surgió del deseo de producir un arma mecánicamente superior. El proceso que abocó en su invención podría guardar relación con la difusión del transporte a caballo durante el tercer milenio a.C. en Asia, cuando se extendió el uso de carros y caballería en las guerras. Es muy probable que los jinetes asiáticos se decidieran a disponer de un arco más corto y, por ello, aumentaran su eficacia y fiabilidad reforzándolo con otros materiales. La evolución del arco compuesto en Asia reflejaría, probablemente, el diseño de los arcos en América del Norte después del siglo XVI. Los jinetes indios de las llanuras experimentaron con el diseño de arcos encolando tendones a los lomos de sus armas. Andando el tiempo, numerosas tribus eliminaron la madera y la sustituyeron en el vientre del arco por cuerno de ante o musmón, perfeccionamiento que está a un solo paso del arco compuesto auténtico.

De los arcos compuestos más antiguos que nos ha sido dado conocer, merece destacarse el arco angular asiático occidental. Apareció durante el tercer milenio a.C. Cuando está armado, adopta la forma de un triángulo de poca altura y la de un semicírculo, cuando está a la máxima tensión. Ilustraciones de estos arcos aparecen en sellos mesopotámicos, en pinturas de tumbas egipcias y en bajorrelieves asirios; ese manojo de pruebas nos habla de su uso a lo largo de casi dos milenios, desde el 2400 hasta el 600 a.C.

Aparte de representaciones artísticas, los arqueólogos han hallado numerosos ejemplares de arcos compuestos angulares en cámaras funerarias egipcias. En 1922, Howard Carter recuperó el famoso legado de la tumba de Tutankhamen: contenía, entre otras piezas, 32 arcos compuestos angulares, 14 arcos sencillos de madera y 430 flechas, amén de carcajes y estuches para arcos.

Las explicaciones iniciales del funcionamiento del arco compuesto angular no podían justificar su manifiesta aptitud para flexionarse a través del mango. Esto contrastaba rotundamente con el diseño del arco inglés tradicional, que "brincaría", o vibraría, en la mano si no fuera por la empuñadura rígida.

Con nuestra réplica de arco compuesto angular hemos comprobado que la porción central de éste se flexiona sólo en apariencia. De hecho, el codo central es inflexible y la flexión tiene lugar sólo a lo largo de los brazos. Al soltar la cuerda no se produce ningún brinco, con lo que la descarga es suave y precisa. La tensión de estos arcos es muy larga, alcanzando los 101 centímetros con los brazos bajo el máximo esfuerzo, por lo que cabe atribuirles una capacidad notablemente acrecentada con respecto a los arcos sencillos del segundo milenio a.C.

El arco angular prevaleció en Asia occidental hasta finales del siglo VII a.C., cuando los escitas se aliaron con los medos y babilonios para destruir el Imperio Asirio. Los escitas, excelentes jinetes y arqueros, al parecer eran originarios de las estepas de la actual Ucrania oriental. Pueblo nómada, ocuparon vastas zonas de Asiria y dejaron muestras de sus características puntas de flecha de bronce, pequeñas y trilobadas, de 25 a 50 milímetros de longitud, desde China hasta Grecia.

La mayoría de los datos reunidos

acerca de los utensilios de arquería escitas proceden de representaciones artísticas. Asimismo, de los enterramientos de Pazyryk, en la región de Altai oriental del Asia central soviética, se han rescatado estuches, carcajes y astiles. Aunque debamos siempre tratar con cautela las representaciones artísticas, la constancia con que se reproduce el arco escita nos permite sacar algunas conclusiones.

El tipo de diseño corrientemente aludido como arco escita, el "arcus scythicus" de los romanos, lo utilizaron distintos pueblos a lo largo de mucho tiempo. Este arma ya la habían desarrollado totalmente los cimerios, procedentes del norte del Cáucaso, en el siglo IX a.C. Los escitas lo introdujeron más tarde entre los antiguos griegos. Finalmente, el modelo acabó llegando incluso hasta Francia septentrional.

Las ilustraciones contemporáneas de los mismos y las medidas de los estuches desenterrados en el yacimiento de Pazyryk muestran que el arco escita tenía una longitud de unos 120 centímetros. En forma de arco de Cupido, su mango era retrasado y sus brazos, recurvados, estaban rematados por extremos ganchudos. Con esa longitud y configuración, incluidos un acusado recurvamiento en la parte del mango y unos brazos flexibles, dan una tensión de unos 76 centímetros.

Este valor concuerda con la longitud de las flechas recuperadas en los enterramientos de Pazyryk. Algunos expertos opinan que la tensión es de sólo 45 centímetros, basándose en figuras decorativas de vasijas áticas del siglo V a.C. y en el pequeño diámetro que presenta el encastre de la base de las puntas de flecha trilobada. El tamaño menudo de la punta de flecha escita y su encastre sugerían, sin duda, que el astil tenía una longitud y un diámetro reducidos en proporción a aquéllos.

Pero esta inferencia no se compagina con la ventaja principal del arco compuesto: su elevada relación tensión/longitud. Así, un arco de 127 centímetros, relativamente corto, podría tensarse mucho más de lo que sugiere esa longitud. Los artesanos estrechaban probablemente la base de las puntas de flecha para que encajaran en un astil abombado. Este contorno, vigente hasta la época medieval, era ahusado con su punto más ancho en el centro del astil. Lo mismo que el arco angular, el escita podría haber sido completamente flexible. Sus brazos carecían de la rigidez inherente que lograban los arcos compuestos al incorporar placas de cuerno o hueso al mango o a los extremos, u orejas, de los brazos.

Suele ocurrir que las armas evolucionen al par que los esfuerzos por lograr una protección fiable contra ellas. En el siglo III a.C. los vecinos orientales de los escitas desarrollaron nuevas técnicas guerreras. Los sármatas, nombre de esos fronterizos, cubrieron con armaduras sus jinetes y monturas y los entrenaron para combatir en formación cerrada. Las duras armaduras corporales demandaban un arco capaz de impulsar una flecha dotada de una pesada punta de hierro a mayor velocidad e impacto.

Los pueblos nómadas del Asia central, pensemos en hunos y ávaros, proporcionaron los medios para traspasar las armaduras. Reforzaron la rigidez de las orejas de los arcos y confirieron a las mismas un contorno ganchudo. Formaban así, en el extremo de cada brazo, una palanca compuesta. Palancas que permitían al arquero flexionar un brazo más pesado con menos esfuerzo; la inclinación con que la oreja se curva hacia el lomo del arco le induce a actuar cual si en el extremo de cada brazo se hubiera fijado una polea.

Conforme el arquero tensa el arco, la "polea" se desenrolla y se alarga la cuerda del arco. Al soltar ésta, las orejas retornan hacia adelante y se acorta la cuerda; la flecha recibe entonces una mayor aceleración. Esta invención precede en muchos siglos al moderno arco mixto, en el que se emplea un sistema de poleas para lograr un efecto similar aunque más pronunciado.

Hacia el siglo XVII, turcos otomanos y tribus turcas de Irán habían modificado el diseño básico del arco compuesto de otros nómadas asiáticos. Los constructores turcos comenzaron a experimentar con arcos acortados hasta unos 111 o 116 centímetros. Prescindieron del mango retrasado y de las placas de hueso y cuerno que se utilizaron en las orejas de las armas anteriores. Crearon de esa forma un arco que se curvaba airosamente, a cada lado de un mango rígido, hacia unos extremos levemente ganchudos.

Esos arcos cortos y poderosos poseían una tensión más larga. Su potencia estaba comprendida entre 36 y



2. ARMADOS DE ARCOS COMPUESTOS, personajes persas acorralan a quienes son hombres lobos. Se trata de una escena mítica de las varias que figuran en un libro del siglo XVIII que describe la vida de un príncipe persa. Los nombres de éste, del artista y del escritor se desconocen.

más de 45 kilogramos, equivalente a la del arco inglés que es el doble de largo. Armada del arco turco, la caballería otomana se reveló formidable, conquistando Europa oriental durante la Edad Media. La invención de la pólvora y el mosquete provocó la caída gradual del uso bélico del arco turco; aunque siguió disfrutando de favor como arma deportiva, especialmente para tiro al vuelo. Acaso el disparo mejor del que se tiene noticia sea el que realizó en 1798 el sultán otomano Selim III, quien lanzó dos flechas al vuelo a 889 metros, hazaña jamás igualada con utensilios de arquería construidos según los patrones tradicionales.

Los útiles de arquería, lo mismo que cualquier otro artefacto de consumo, reflejan la naturaleza de los materiales al alcance y de las necesidades de los individuos. Por citar un ejemplo, los arcos militares compuestos, grandes y pesados, que construyeron los tártaros de Crimea en el siglo XVII convierten en juguetes de niños los arcos de los modernos bosquímanos de Kalahari. Los arcos bosquímanos, no obstante, cumplen magníficamente su labor en su propio contexto; es decir, pueden catapultar una flecha pequeña no emplumada que inocula veneno bajo la piel del animal que alcanza.

En pocas palabras, el rendimiento del modelo puede sólo calibrarse por su aptitud para actuar con éxito en su contexto de aparición. Durante milenios, el arco se mantuvo como el más eficaz instrumento lanzador de proyectiles de que disponían los cazadores y guerreros. Su supremacía no comenzó a declinar hasta el siglo XVI con la invención y difusión de un arma todavía más poderosa: el arma de fuego.

BIBLIOGRAFIA COMPLEMENTARIA

NEOLITHIC BOWS FROM SOMERSET, ENGLAND, AND THE PREHISTORY OF ARCHERY IN NORTH-WEST EUROPE. J.G.D. Clark en *Proceedings of the Prehistoric Society*, vol. 29, págs. 50-98; diciembre de 1963.

THE BOW: SOME NOTES ON ITS ORIGIN AND DEVELOPMENT. Gad Rausing. Bonn, Rudolf Habelt, y Lund, CWK Gleerups, 1967.

TURKISH ARCHERY AND THE COMPOSITE BOW. Tercera edición. Paul E. Klopsteg. Manchester, Simon Archery Foundation, 1987.

EXPERIMENTAL ARCHERY: PROJECTILE VELOCITIES AND COMPARISON OF BOW PERFORMANCES. C. A. Bergman, E. McEwen y R. Miller en *Antiquity*, vol. 62, n.º 237, págs. 658-670; diciembre de 1988.

Cirugía con láser

El láser, herramienta precisa, poderosa y delicada, está aumentando su importancia en el campo de la medicina. Estas cuchillas de luz permiten manipular lo mismo una célula que un órgano entero

Michael W. Berns

El láser es un eficaz escalpelo. No saja cuanto le sale al paso. Actúa de forma tan selectiva, que lo mismo penetra en el interior de una célula, que se adentra en un órgano dejando intacto lo demás, proeza inalcanzable por cualquier bisturí.

El refinamiento de su precisión en los últimos 30 años ha permitido ampliar las aplicaciones médicas del láser, desde sus comienzos en que se aprovechaba el calor generado por el haz lasérico para destruir el tejido. Los efectos térmicos siguen hoy constituyendo la razón principal de su uso médico, pero empiezan a abrirse camino en el diagnóstico y la terapia otros efectos ajenos al calor. Además de calentar los tejidos, los fotones del haz del láser desencadenan reacciones químicas, rompen los enlaces atómicos que mantienen unidas a las moléculas y crean ondas de choque.

Podríamos citar muchas y dispares aplicaciones biomédicas del láser; por citar algunas: limpieza de la luz de las arterias, disgregación de cálculos renales, eliminación de cataratas e incluso alteración del material genético. Con el láser extraemos información del funcionamiento interno de las células, cuyo interés no se ciñe a la biología, sino que redunda en beneficio de la medicina.

Quizá por ese paraíso prometido, la técnica del láser se ha presentado en un celofán de ensueño e hipérbole. El futuro realista de la cirugía lasérica descansa en una mejor comprensión de los mecanismos básicos, físicos y químicos, que involucran la interacción entre la luz y los órganos u orgánulos. Conocimiento que impondrá un criterio para establecer cuándo está o no indicado el empleo del láser. El saber reconocer que un escalpelo o una cuchilla de electrocauterización de 50.000 pesetas resultan, en determinadas circunstancias, más idóneos y menos peligrosos que un ingenio lasérico de un millón de pesetas constituye un requisito fundamental para que esta nueva técnica proceda abriéndose camino en el campo de la medicina.

MICHAEL W. BERNS ocupa la cátedra Arnold and Mabel Beckman de la Universidad de California en Irvine, centro superior con el que tiene contrato desde 1972. Se formó en la Universidad de Cornell. Berns y el industrial Arnold O. Beckman fundaron, en 1982, el Instituto Beckman del Láser, donde trabajan clínicos, biólogos, ingenieros y físicos en la investigación y aplicación de esa técnica. Berns dedica este artículo a Arnold O. Beckman en la celebración de su 91 cumpleaños.

La idea de utilizar la luz con fines quirúrgicos precedió, en varios años, al descubrimiento del láser. En 1946, Gerd Meyer-Schwickerath, médico alemán, se servía de la luz solar para tratar desprendimientos de retina y destruir tumores oculares en sus pacientes. En 1961, sólo un año después de que Theodore H. Maiman construyera el primer láser en los Laboratorios de Investigación Aérea de Hughes, Milton Zaret, de la facultad de medicina de la Universidad de Nueva York, lo manejaba ya para provocar lesiones oculares en animales. Los ensayos con humanos comenzaron dos años después, cuando Chris Zweng, de la Fundación de Investigación Médica de Palo Alto, en California, trató mediante láser las enfermedades retinianas de sus pacientes. No tardó en aceptarse como una pieza más del armamentario quirúrgico habitual en oftalmología.

Los primeros cirujanos recurrían a ese haz luminoso por el calor intenso que generaba. Esa propiedad sigue explotándose todavía, merced a la especificidad de su efecto destructor y al control preciso que permite tener de su operación. Si la longitud de onda de la luz del láser coincide con la banda de absorción de la zona diana, esta zona absorberá la luz del láser y sufrirá la agresión consiguiente.

Demos un ejemplo. La melanina de la retina es un pigmento de color marrón oscuro que absorbe el haz verde del láser de argón. Ello nos dice que el láser de argón puede destruir regiones específicas de la retina sin dañar otras zonas del ojo, que absorben luz de diferentes longitudes de onda. Mediante este procedimiento se puede tratar eficazmente la retinopatía diabética, una enfermedad degenerativa que causa buena parte de las cegueras adquiridas.

Las manchas rojas de nacimiento absorben también el haz del láser de argón, que será azul o verde a tenor de su longitud de onda. La luz destruye los cientos de vasos sanguíneos que se encuentran debajo mismo de la capa exterior de la piel y le confieren el color característico. Aunque en este caso la cirugía con láser es preferible a la incisión e injerto de la piel, la técnica tiene sus inconvenientes. El calor generado por el haz podría extenderse a otras zonas adyacentes a los vasos sanguíneos anómalos y producir cicatrices o pérdidas de pigmentación.

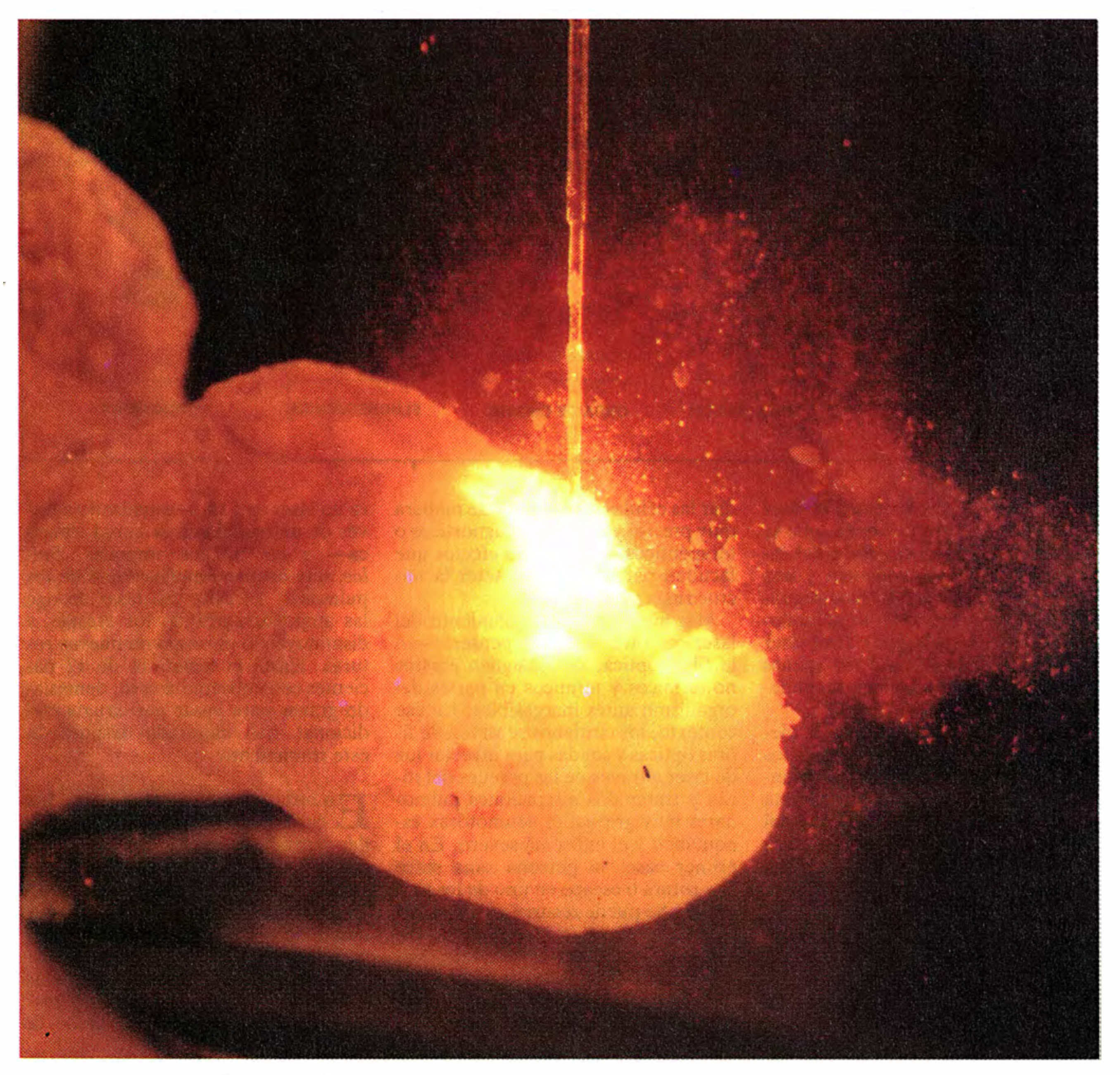
El evitar esos efectos secundarios constituyó un hito en la cirugía lasérica. En 1983, R. Rox Anderson y John A. Parrish, de la Universidad de Harvard, sugirieron que exposiciones cortas, inferiores a una milésima de segundo, a luz intensa destruirían la zona de absorción sin dañar el tejido adyacente. El proceso de absorción de la energía y la subsiguiente disipación del calor requerían, argumentaban, menos tiempo que la transferencia de ese calor a las zonas contiguas. Por tanto, la destrucción selectiva de las zonas diana pigmentadas tendría dos requisitos: absorción de luz preferente y pulsación luminosa suficientemente corta.

Se demostró que esta teoría era cierta. La fototermólisis selectiva, nombre que recibe la técnica, ha per-

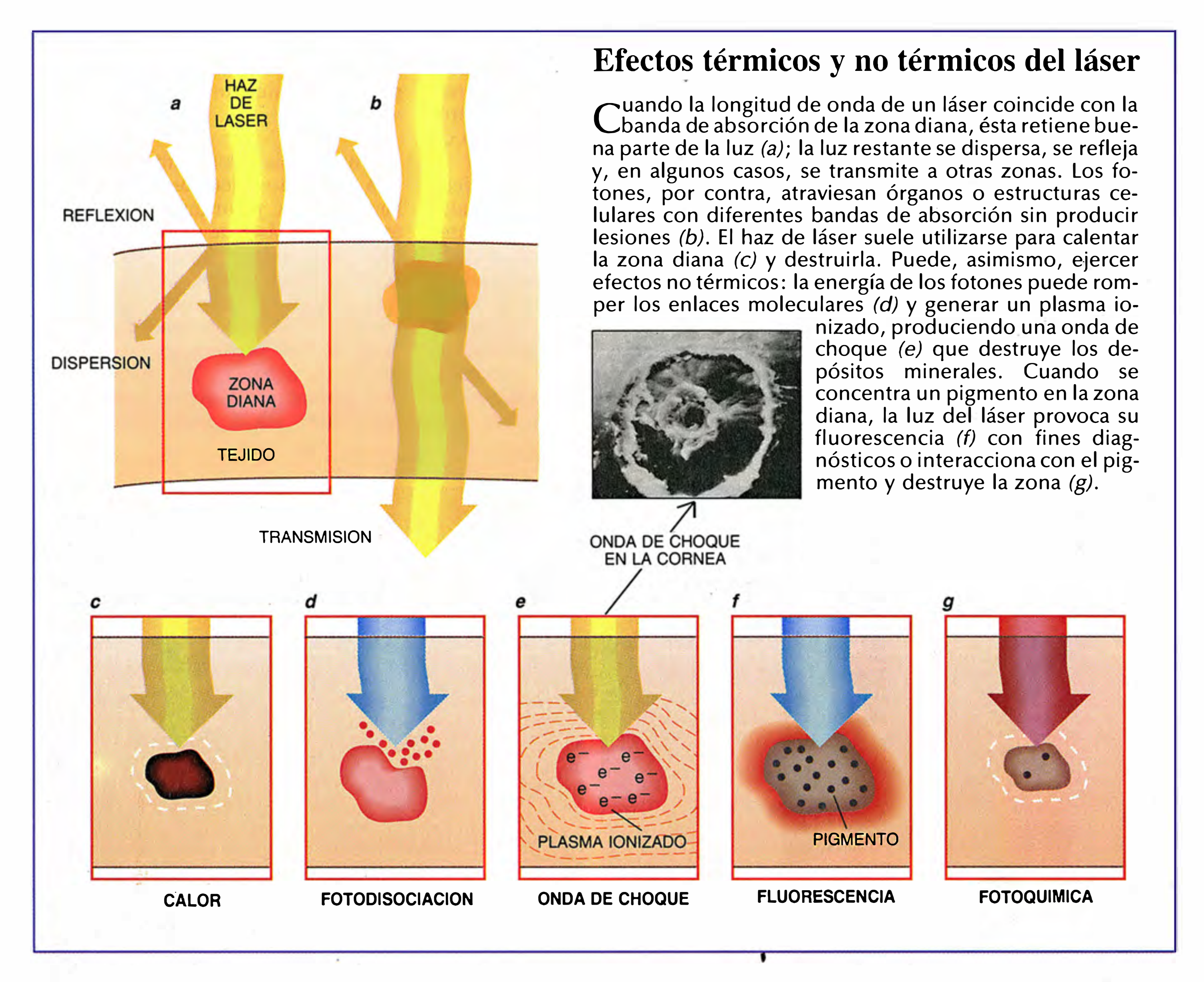
mitido mejorar notablemente el tratamiento de las manchas rojas de la piel. También se ha mostrado eficaz para eliminar tatuajes. Se puede evitar la cicatrización emitiendo el haz de láser en pulsos cortos, en vez de emitirlo de forma continua o en pulsos largos, que duran apenas la cuarta parte de un segundo. (Lo ideal sería que admitiesen un uso continuo para propagar los efectos térmicos allende la zona de absorción o un empleo en pulsos cortos para circunscribir la destrucción a la zona diana.)

Ahora bien, en determinadas circunstancias, la propia extensión de las lesiones que produce el calentamiento mayor y más prolongado del tejido puede resultar una ventaja. Pensemos en el cirujano que se propone destruir una zona dañada del hígado sin producir hemorragias importantes; o en el ginecólogo que desea extirpar un tumor cervical maligno en fase inicial y utilizar simultáneamente el calor para obturar los capilares adyacentes que contribuyen al sangrado. En ambos casos, la exposición prolongada al láser de onda continua (a diferencia del láser de pulsos cortos) reduce la hemorragia gracias a que el calor se extiende a los capilares cercanos. Para estas situaciones, podemos recurrir a un láser de CO_2 con una longitud de onda de 10,6 micrometros, ya que es absorbido por el componente dominante en los tejidos: el agua.

Aunque para determinados procedimientos médicos se exigirá un haz continuo, y por tanto con efectos térmicos, el láser de pulsos sirve también para destruir tejido. Mi colaborador J. Stuart Nelson ha demostrado que el láser de erbio-itrio-aluminio-granate (YAG), con una longitud de onda de 2,9 micrometros y una duración de pulso de 200 microsegundos, elimina con limpieza tejido óseo calcificado. En la parte opuesta del



1. CALCULO EN EL TRACTO URINARIO, uno de los muchos depósitos calcificados, con los formados en el riñón y la vesícula biliar, que el láser destruye. En el ejemplo de la fotografía tomada por Peter Hering, la energía del haz del láser, liberada a través de una fibra óptica de 400 micrometros, genera un plasma ionizado. Este plasma produce, a su vez, ondas de choque que destruyen el cálculo.



Efectos térmicos y no térmicos del láser

Cuando la longitud de onda de un láser coincide con la banda de absorción de la zona diana, ésta retiene buena parte de la luz *(a)*; la luz restante se dispersa, se refleja y, en algunos casos, se transmite a otras zonas. Los fotones, por contra, atraviesan órganos o estructuras celulares con diferentes bandas de absorción sin producir lesiones *(b)*. El haz de láser suele utilizarse para calentar la zona diana *(c)* y destruirla. Puede, asimismo, ejercer efectos no térmicos: la energía de los fotones puede romper los enlaces moleculares *(d)* y generar un plasma ionizado, produciendo una onda de choque *(e)* que destruye los depósitos minerales. Cuando se concentra un pigmento en la zona diana, la luz del láser provoca su fluorescencia *(f)* con fines diagnósticos o interacciona con el pigmento y destruye la zona *(g)*.

espectro visible se encuentra el láser excímero de cloruro de xenón, que se sitúa en la región ultravioleta del espectro con 0,308 micrometros y una duración de pulso de 10 nanosegundos (un nanosegundo es la milmillonésima parte del segundo). Este láser puede vaporizar el hueso sin apenas producir lesiones térmicas secundarias.

Aunque estos dos tipos de láser vienen a producir idénticos efectos en los tejidos, operan de manera muy distinta. La energía del fotón ultravioleta decuplica la del fotón procedente del láser de erbio YAG, energía que se destina probablemente a la ruptura de los enlaces moleculares en el tejido diana a través de un proceso no térmico llamado fotodisociación molecular.

Cuando el tejido y sus células absorben la intensa luz del láser, la energía debe disiparse. Esta pérdida adquiere distintas formas: calor (ya explicado), fotodisociación, ondas de choque, reacciones químicas o fluorescencia. Los médicos utilizan todos estos efectos para manipular o estudiar las células y los tejidos de manera muy precisa y con fines diagnósticos o terapéuticos. Abanico de efectos que facultan para mejor acometer la microcirugía subcelular.

Más aún. Con el acoplamiento del láser con otras técnicas, pensemos en la fibra óptica, se consiguen efectos no térmicos y térmicos en partes del organismo antes inaccesibles. En ese contexto, los cirujanos se sirven de fibras ópticas y sondas para inyectar luz de láser a través de las paredes del tórax y tratar dos alteraciones pulmonares muy graves: el neumotórax espontáneo y el enfisema severo. En el primer caso, la persona sana sufre una rotura o escape en uno de los pulmones. Se puede recurrir al láser para cerrar el escape, tal y como descubrieron mis colaboradores Akio Wakabayashi y Matthew Brenner, de la Universidad de California en Irvine. No se necesita, pues, la intervención quirúrgica tradicional.

El mismo procedimiento viene en nuestro auxilio para tratar el enfisema, que, en distintos grados, afecta a muchos millones de personas. Se aplica un láser de CO_2 —canalizado a través de un orificio en la pared torácica— a las frágiles ampollas, *bullae*, que ocupan amplias zonas de los pulmones. El calor del láser encoge las ampollas cerrando los escapes y disminuyendo el riesgo de nuevas roturas. Hasta la fecha, 11 de 12 pacientes con enfisema severo, demasiado graves para pasar por cirugía tradicional, han mejorado después de este tratamiento.

En otra aplicación de la técnica, los cardiólogos y los radiólogos pueden introducir en los vasos sanguíneos fibras de 400 micrometros —o un haz flexible de hasta 400 fibras de 50 micrometros— y llegar a una zona obstruida en el sistema circulatorio periférico o coronario. A través de la fibra óptica, se transmite la luz del láser que acaba con el bloqueo y restablece la circulación normal de la sangre.

La angioplastia con láser, nombre que recibe esta técnica, se hacía originariamente con haces de láser térmicos en combinación con la angio-

plastia de globo. Correspondía a la sonda de láser abrir paso a través de un vaso sanguíneo total o parcialmente ocluido. Se introducía luego el globo y se hinchaba, dilatando más la luz del vaso.

Pero el enfoque térmico no resultaba idóneo para eliminar los depósitos calcificados que caracterizan a la arterioesclerosis. Mis colaboradores Jonathan M. Tobis y Walter L. Henry, de Irvine, descubrieron que, con cierta frecuencia, los depósitos duros desviaban las fibras y el haz perforaba el vaso. Igualmente, tal y como veíamos en el caso de las manchas rojas de nacimiento, el calor puede lesionar más allá de la zona diana. En algunos casos daña la pared vascular normal adyacente.

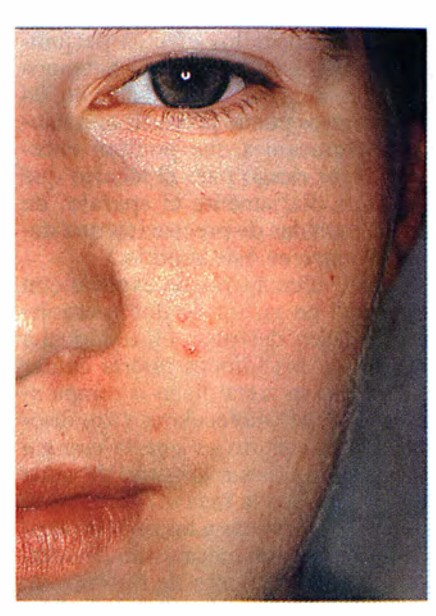
Por contra, el láser excímero de 0,308 micrometros parecía ideal para la angioplastia, con o sin globos. Pares excitados, o dímeros, de átomos de argón y flúor, o de átomos de xenón y cloro, liberan energía, generando el haz de láser excímero (excímero es término convencional para aludir a los "dímeros excitados"). La fibra óptica flexible de cuarzo transmite eficazmente la longitud de onda del excímero de 0,308 micrometros. James S. Forrester, Frank Litvack y Warren S. Grundfest, del Hospital de los Cedros del Sinaí en Los Angeles, demostrarían que el láser excímero podía restaurar la circulación coronaria.

Las angioplastias coronarias realizadas con la ayuda del láser se han multiplicado, con una tasa de complicaciones pareja a las que se siguen de la aplicación de las técnicas de angioplastia tradicionales. La cifra actual de reoclusiones viene a ser la misma que la de la angioplastia con globo, pero es de esperar que, con el paso del tiempo, una técnica basada en la destrucción y desintegración de la placa dura será más eficaz que el mero ensanchamiento del vaso con la ayuda de un globo. Cualquiera de las dos técnicas constituye una vía alternativa a la cirugía de derivación coronaria, cara y arriesgada, que precisa, además, de un largo período de hospitalización.

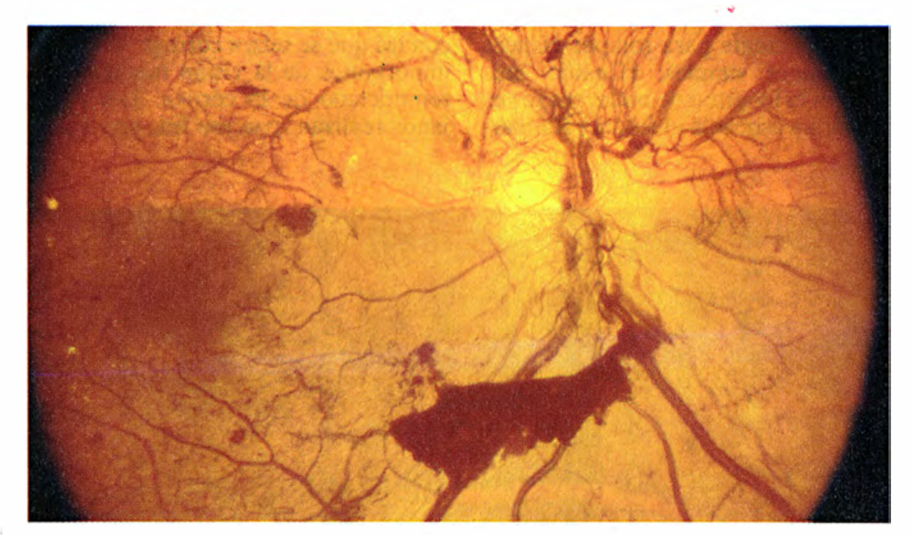
De momento, la angioplastia coronaria con láser sigue en fase experimental. Para una aplicación generalizada de esta técnica habrá que esperar que se consiga el dominio de los equipos de láser y de representación de imágenes. La combinación de imágenes digitales y técnica de ultrasonidos ofrecerá, andando el tiempo, métodos fiables de representación visual directa para el gobierno del láser.

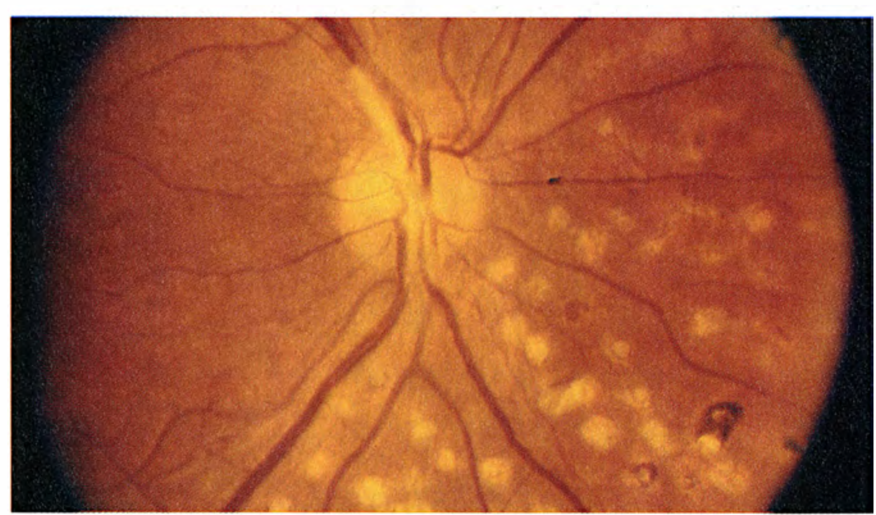
Amén de los técnicos hay que salvar obstáculos biológicos. No acaba-





2. MANCHAS ROJAS DE NACIMIENTO y su terapia lasérica. Los vasos sanguíneos supernumerarios que yacen debajo mismo de la capa exterior de la piel (*izquierda*) absorben la luz de láser amarilla, que destruye los vasos (*derecha*). Como el haz se libera en pulsos breves, los tejidos adyacentes no sufren daños.





3. RETINOPATIA DIABETICA, causa de ceguera adquirida. Se caracteriza por un crecimiento excesivo de vasos sanguíneos en la retina y algo de hemorragia (*arriba*). Podemos focalizar un láser de argón hacia los vasos sanguíneos para detener la hemorragia y hacia otras zonas de la retina donde el pigmento marrón melanina absorbe la luz del láser. Este procedimiento causa quemaduras circulares (*abajo*) que retardan el crecimiento de nuevos vasos.

mos de conocer los efectos perjudiciales del láser excímero. Sabemos que una longitud de onda de 0,308 micrometros puede producir mutaciones. (Mutaciones que podrían encerrar mayor riesgo para el técnico que manipula diariamente el aparato, de ahí las medidas de precaución que deben tomarse en este sentido.)

Nadie, además, puede explicar con seguridad en virtud de qué mecanismo el láser destruye el tejido. La energía de los fotones podría actuar calentando la zona diana o rompiendo los enlaces moleculares (fotodisociación). Lo cierto es que la energía de un fotón de esta longitud de onda se encuentra en la raya divisoria entre los efectos térmicos y los no térmicos.

Sí está claro el mecanismo destructor del láser excímero de fluoruro de argón: rompe los enlaces. Los fotones procedentes de este láser, con una longitud de onda ultravioleta de 0,193 micrometros, poseen casi un tercio más de energía que los de longitud de onda de 0,308 micrometros. Esa energía es suficiente para romper enlaces moleculares simples. (Cuanto más corta es la longitud de onda, mayor es la energía asociada al fotón.) Rangaswamy Srinivasan, de la empresa IBM, pionero en el uso del láser excímero de 0,193 micrometros para grabar sobre películas de polímero, demostró sin ambages esa propiedad.

La precisión de la fotodisociación abrió la puerta para la cirugía ocular con láser. En 1983, Stephen L. Trokel, de la Universidad de Columbia, demostró, en colaboración con Srinivasan, que la ruptura de los enlaces facilitaba la eliminación de pequeñas cantidades de tejido de la córnea del ojo. El grupo de Trokel determinó la relación entre la energía liberada por el láser y la eliminación del tejido ocular: cada pulso de 10 nanosegundos eliminaba 0,2 micrometros de tejido (aproximadamente una milésima del diámetro de una célula). La eliminación de tejido era precisa y limpia; además, no se veía que el tejido adyacente sufriera daños.

Los descubrimientos de Trokel condujeron a la aplicación del láser excímero de 0,193 micrometros a las técnicas de escultura de la córnea, con las que se busca corregir defectos importantes de la visión mediante el remodelado de la córnea. Los cirujanos realizan, con ese fin, pequeños cortes lineales u optan por eliminar zonas superficiales amplias. Si la fortuna acompaña, podrían corregirse así miopías, hipermetropías y astigmatismos, entre otros defectos de la visión. Los cirujanos oculares también utilizan esta técnica para eliminar cicatrices de la córnea y crecimientos anómalos de tejido. Cada una de estas aplicaciones del láser excímero, en fase de controvertida experimentación, se está valorando actualmente en ensayos clínicos con seres humanos.

Antes de que la escultura de la córnea se convierta en procedimiento rutinario conviene tener respuesta a varias preguntas. No sabemos si la cirugía produce mutaciones, si la córnea se recupera adecuadamente, cuáles son los efectos a largo plazo de las lesiones en la córnea y si los equipos poseen garantía suficiente de precisión y seguridad. El coste de estos sistemas de láser —cifrado en decenas de millones de pesetas— es también elevado.



4. PIGMENTO FLUORESCENTE que se retiene de forma selectiva en la piel del ratón, por más que aquí sólo sea visible en la cola y las orejas. Cuando se ilumina al animal con luz azul, el pigmento emite fluorescencia. De manera similar, el pigmento que se acumula en los tejidos cancerosos emite fluorescencia en cuanto se le expone a cierta luz de láser; de ahí su interés en diagnóstico.

Además de romper los enlaces moleculares, el láser induce ondas de choque, otro prometedor efecto clínico no térmico. Las ondas de choque encuentran razón de ser en la técnica quirúrgica de eliminación de cataratas secundarias, así se llaman las que se forman en la membrana posterior de los cristalinos artificiales y que padecen el 30 por ciento de los enfermos que tienen estos implantes.

Hasta 1980 sólo había una manera de tratar las cataratas secundarias: la incisión quirúrgica de la membrana posterior, con obligada anestesia general. Danièle Aron-Rosa, de la Universidad de París, y Franz Fankhauser, de la Universidad de Berna, demostraron que podía enfocarse un láser infrarrojo de ondas cortas hacia esta membrana opaca posterior, o en su contorno, produciéndose su desgarro por una onda de choque. Emplearon un láser YAG de neodimio, que puede generar pulsos del orden del nanosegundo o picosegundos (billonésima parte de un segundo) y funcionar a una longitud de onda de 1,06 micrometros.

Después del tratamiento con el láser, la visión del paciente mejora casi de inmediato. En Estados Unidos se llevan a cabo anualmente más de 200.000 operaciones de este tipo, llamadas capsulotomías posteriores. A diferencia de la técnica quirúrgica anterior, que tenía un coste de unas 200.000 pesetas, la cirugía con láser cuesta alrededor de 100.000 pesetas y no requiere anestesia general ni hospitalización.

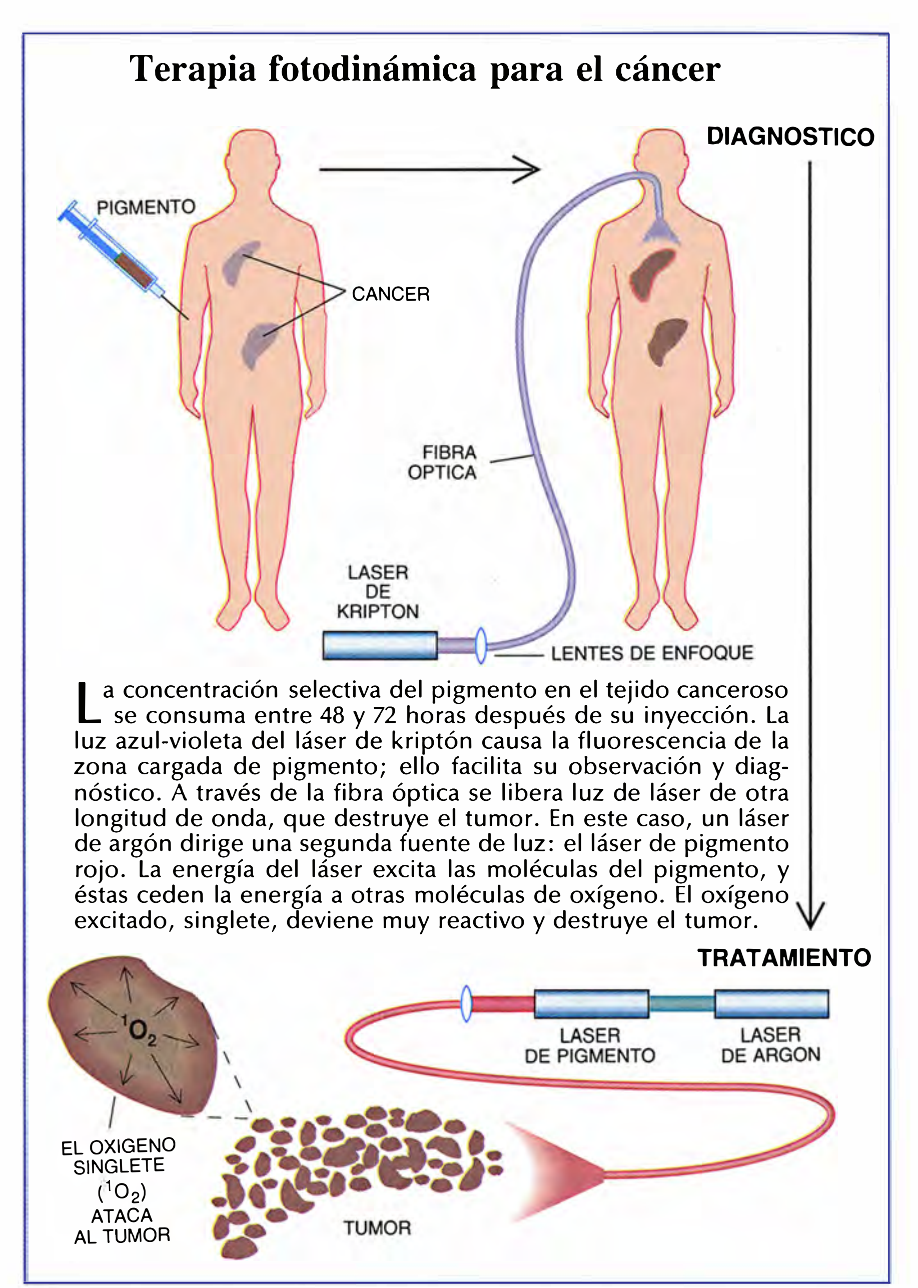
Ese recurso al láser de ondas cortas nos lo posibilita la intensidad de un haz apretadamente focalizado: se liberan milijoules de energía en un intervalo temporal de 10^{-12} a 10^{-9} segundos y en una superficie de 25 a 50 micrometros de diámetro. El haz atraviesa la córnea externa y el cristalino artificial, estructuras que son transparentes a la longitud de onda de 1,06 micrometros, para concentrarse de forma selectiva en la sustancia vítrea, o gelatinosa, adyacente a la catarata secundaria. (En ciertos procedimientos en los que se aplica el láser, se puede alcanzar tal intensidad de haz que se lesione sólo el punto focalizado. Quiere ello decir que ni las demás estructuras ni tejidos sufrirán daño alguno a su paso.)

En el caso de la capsulotomía posterior, los fotones atraviesan la córnea y el cristalino. Sin embargo, en el punto focal cercano a la catarata secundaria, la densidad de fotones es tan elevada que los átomos pierden electrones en un proceso denominado cesión óptica o ionización. Los electrones forman una nube gaseosa altamente excitada, o plasma, que atrapa los restantes fotones procedentes del láser. En virtud de esta absorción, aumenta la temperatura en el punto focal, que no tarda en elevarse hasta decenas de miles de grados Celsius. El plasma se expande rápidamente y produce una onda de choque que se propaga en todas direcciones. La onda de choque rompe la catarata secundaria. Sin embargo, el plasma no se crea hasta que los fotones alcanzan cierto umbral de intensidad. Una vez sucede esto, la fuerza de la onda de choque se hace proporcional a la cantidad de energía absorbida.

Los cálculos del riñón, uréter y vesícula biliar podrían someterse también a la acción de las ondas de choque. La propagación de haces de pulsos cortos, a través de la uretra o el uréter, con la ayuda de la fibra óptica rompería los cálculos. Las piedras en la vesícula biliar pueden fragmentarse mediante un procedimiento nuevo y bastante similar. El cirujano inserta un endoscopio de fibra óptica, llamado laparoscopio, a través de un pequeño orificio practicado en el abdomen del paciente. Utilizando una cuchilla de electrocauterización o un haz de láser térmico de onda continua, se desprende la vesícula enferma del hígado para extirparla a través de la pequeña incisión abdominal.

Aunque no sea imprescindible para la disección, el láser resulta idóneo cuando la vesícula biliar tiene cálculos. Por culpa de estos depósitos duros, el cirujano no puede extirpar la vesícula a través de la incisión, dificultad que orilla con la introducción directa de un láser de pulsos en la vesícula biliar a través del laparoscopio. Las ondas de choque fragmentan los cálculos y la vesícula se puede extirpar fácilmente. La operación resulta mucho menos traumática para el paciente que la cirugía mayor.

Terapia fotodinámica para el cáncer



La concentración selectiva del pigmento en el tejido canceroso se consuma entre 48 y 72 horas después de su inyección. La luz azul-violeta del láser de kriptón causa la fluorescencia de la zona cargada de pigmento; ello facilita su observación y diagnóstico. A través de la fibra óptica se libera luz de láser de otra longitud de onda, que destruye el tumor. En este caso, un láser de argón dirige una segunda fuente de luz: el láser de pigmento rojo. La energía del láser excita las moléculas del pigmento, y éstas ceden la energía a otras moléculas de oxígeno. El oxígeno excitado, singlete, deviene muy reactivo y destruye el tumor.

La destrucción inmediata de una catarata o la fragmentación de un cálculo renal son aplicaciones espectaculares del láser. Pero consiguen efectos más sutiles. El láser puede inducir reacciones químicas, igual que el sol promueve la fotosíntesis. Lo cierto es que la fotoquímica del láser podría explotarse muy pronto para tratar el cáncer.

Una vez más, esta aplicación curativa de la luz constituye el cumplimiento actual de observaciones realizadas en el pasado. Se observó, a principios de siglo, que los tejidos cancerosos concentraban pigmentos corporales, en particular la porfirina, el pigmento rojizo de la sangre. Hasta los años setenta no se sacaría partido de esa observación. El grupo de Thomas J. Dougherty, del Instituto Roswell Park Memorial en Buffalo, Nueva York, demostró que los animales inyectados con porfirina retenían altas concentraciones de este pigmento en sus tumores pasadas entre 48 y 72 horas. Cuando se les aplicaba un haz de luz con la longitud de onda que coincidía con la banda de absorción de la porfirina, se detenía el crecimiento y, en ocasiones, el tumor desaparecía.

Sabemos hoy que esa destrucción

del tumor mediada por la presencia del pigmento se produce merced a la generación de un citotóxico, una molécula de oxígeno excitada, llamada oxígeno singlete. (El término "singlete" denuncia el estado de espín de la molécula de oxígeno.) La transferencia de energía de la molécula excitada de porfirina al oxígeno crea el oxígeno singlete excitado. En este estado, el oxígeno es altamente reactivo y, por tanto, tóxico. El oxígeno singlete ataca la membrana celular externa, así como a muchas de las estructuras del interior celular limitadas por membranas, como las mitocondrias y los lisosomas. Se destruye, en definitiva, el tumor sin dañar el tejido normal.

Aunque los primeros estudios esporádicos fueron alentadores, el papel que la sensibilización del pigmento, o terapia fotodinámica, desempeñará en oncología está todavía por determinar. Es necesario llevar a cabo estudios de supervivencia a largo plazo, amén de realizar comparaciones con los otros tratamientos ya establecidos, como la radioterapia, la quimioterapia y la cirugía.

Para algunas instituciones estatales, así la norteamericana Administración de Drogas y Alimentos, la terapia fotodinámica aplicada al cáncer se encuentra en fase experimental. Hubo en un comienzo un grave problema: los enfermos se hacían muy sensibles a la exposición a la luz intensa durante un período que se prolongaba hasta tres meses después de la administración del pigmento. Hoy se trabaja en el desarrollo de mejores técnicas de aplicación y de pigmentos nuevos que no comporten fotosensibilidad residual. Se está comprobando, en efecto, la evolución de este tratamiento aplicado al cáncer de pulmón, de vejiga y de esófago; en los ensayos clínicos multicéntricos intervienen hospitales de Estados Unidos y Canadá.

El potencial diagnóstico de esta técnica nos tiene en vilo. Muchos pigmentos que absorben energía tornan a emitir la luz en forma de fluorescencia, descubriéndonos la presencia de zonas cancerosas. En los primeros estudios realizados en Estados Unidos y Japón, se utilizó la fluorescencia para detectar tumores de pulmón y de vejiga en fases iniciales. Mediante el uso de fibra óptica y detectores de luz, podríase diagnosticar la presencia de pequeños tumores en zonas recónditas del cuerpo. En última instancia, podríamos administrarle al paciente un "revoltillo" de pigmentos, unos para el diagnóstico y otros para la destrucción del tumor instada por la luz.

Es cierto que la salud del cuerpo humano y sus órganos constituye el motivo principal de investigación en las aplicaciones del láser a la medicina. Pero no es el único. Otros objetivos menores merecen también su debida atención. El láser nos ofrece una técnica inédita para el estudio singular de las células y los procesos fotofísicos, que aportan las bases para la aplicación del láser en medicina. En esa línea de investigación venimos laborando en mi laboratorio durante 20 años. El conocimiento de los mecanismos de interacción entre la luz y los orgánulos celulares ha facilitado el estudio de la estructura y las funciones básicas de la célula. La inmensa mayoría de los mecanismos fotofísicos descritos en este artículo pueden aplicarse y estudiarse en el dominio celular y subcelular.

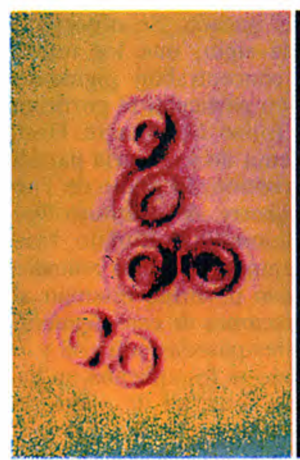
Por ejemplo, se pueden incorporar pigmentos fotoactivos a un segmento de la molécula de ADN. Se puede luego inactivar esa zona mediante la exposición al haz azul-verdoso del láser de argón, produciéndose una supresión genética selectiva. De forma alternativa, un láser ultravioleta de pulsos puede generar un microplasma en la membrana exterior de la célula, produciendo un pequeño orificio que se abrirá lo suficiente para permitir que se introduzca en la célula una cadena de ADN, con genes específicos. Este ADN puede insertarse en la dotación genética celular mediante un procedimiento 10.000 veces más eficaz que la ingeniería genética al uso. Procedimiento que beneficiaría grandemente a la bioingeniería vegetal, por cuanto resulta dificilísimo introducir ADN extraño, con los métodos bioquímicos conocidos, a través de la pared celular rígida de las células vegetales.

Se trabaja también en una perspectiva más general de la perforación celular, que facilite la penetración de los espermatozoides en el óvulo. Desde un enfoque clínico, podría recurrirse al láser ultravioleta de pulsos, finamente focalizado, para abrir un orificio de 10 micrometros en la capa protectora externa del óvulo, por donde entrara sin dificultad un espermatocito. Esa micromanipulación lasérica del óvulo es viable, como confirma la investigación acometida por Ricardo H. Asch, de Irvine, y Yona Tadir, de la Universidad de Tel Aviv.

En esa misma línea, Aaron Lewis, Neri Laufer y Daniel Palanker, de la Universidad Hebrea de Jerusalén, consiguieron, con un haz de láser excímero de 0,193 micrometros, agujerear la capa externa de un óvulo de conejo. La fecundación se produjo a una velocidad mayor que la normal. El método lasérico superará, quizás, en rapidez y especificidad las demás técnicas de penetración asistida del espermatozoide en el óvulo. Queda, sin embargo, mucho por averiguar sobre el potencial mutagénico y los efectos bioquímicos derivados de la energía del láser.

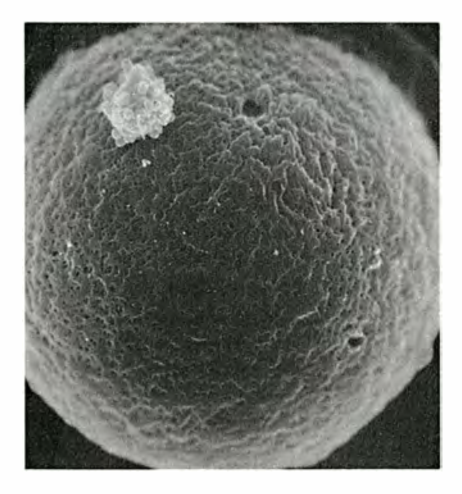
La aplicación reciente más interesante del láser en la cirugía y manipulación celular reside, a buen seguro, en su uso de pinza óptica. Arthur Ashkin, de los Laboratorios AT&T Bell, desentrañó la capacidad del láser para atrapar un objeto en el haz de luz. Se puede utilizar la energía asociada con los fotones para mantener las células quietas, para girarlas o para dirigirlas hacia cualquier dirección.

La facultad de confinar células o sus orgánulos en una trampa óptica, de inmovilizarlos, permite la intervención quirúrgica con un segundo





5. PINZAS OPTICAS, una de las muchas aplicaciones celulares del láser. Los fotones del haz de láser puede mantener quietas las células en un lazo óptico. La luz provoca el movimiento o giro de estos elementos minúsculos. En la composición, el segundo glóbulo rojo gira sobre su eje.



6. CUBIERTA EXTERNA del óvulo de ratón, o zona pelúcida, perforada dos veces por un láser para facilitar la entrada de los espermatocitos.

haz de láser. Podemos fusionar fácilmente las membranas de dos células adyacentes; también, las pinzas pueden fijar un cromosoma mientras un segundo láser secciona un fragmento, que podría aislarse para formar un clon. En el caso precedente de la fecundación del óvulo, podría atraparse el espermatocito y dirigirlo hacia el orificio creado en la capa exterior del óvulo.

Se trate de intervenir en un órgano o en una célula, el láser ofrece a los médicos y a los investigadores una serie de recursos ópticos precisos sin precedentes. En muchas especialidades, la cirugía con láser se ha convertido ya en un procedimiento estándar y rutinario. El tiempo y la investigación clínica tienen la llave para extender su campo de aplicación.

BIBLIOGRAFIA COMPLEMENTARIA

SELECTIVE PHOTOTHERMOLYSIS: PRECISE MICROSURGERY BY SELECTIVE ABSORPTION OF PULSED RADIATION. R. Rox Anderson y John A. Parrish en *Science*, vol. 220, págs. 524-527; 29 de abril de 1983.

EVALUATION AND INSTALLATION OF SURGICAL LASER SYSTEMS. David B. Apfelberg. Springer-Verlag, 1986.

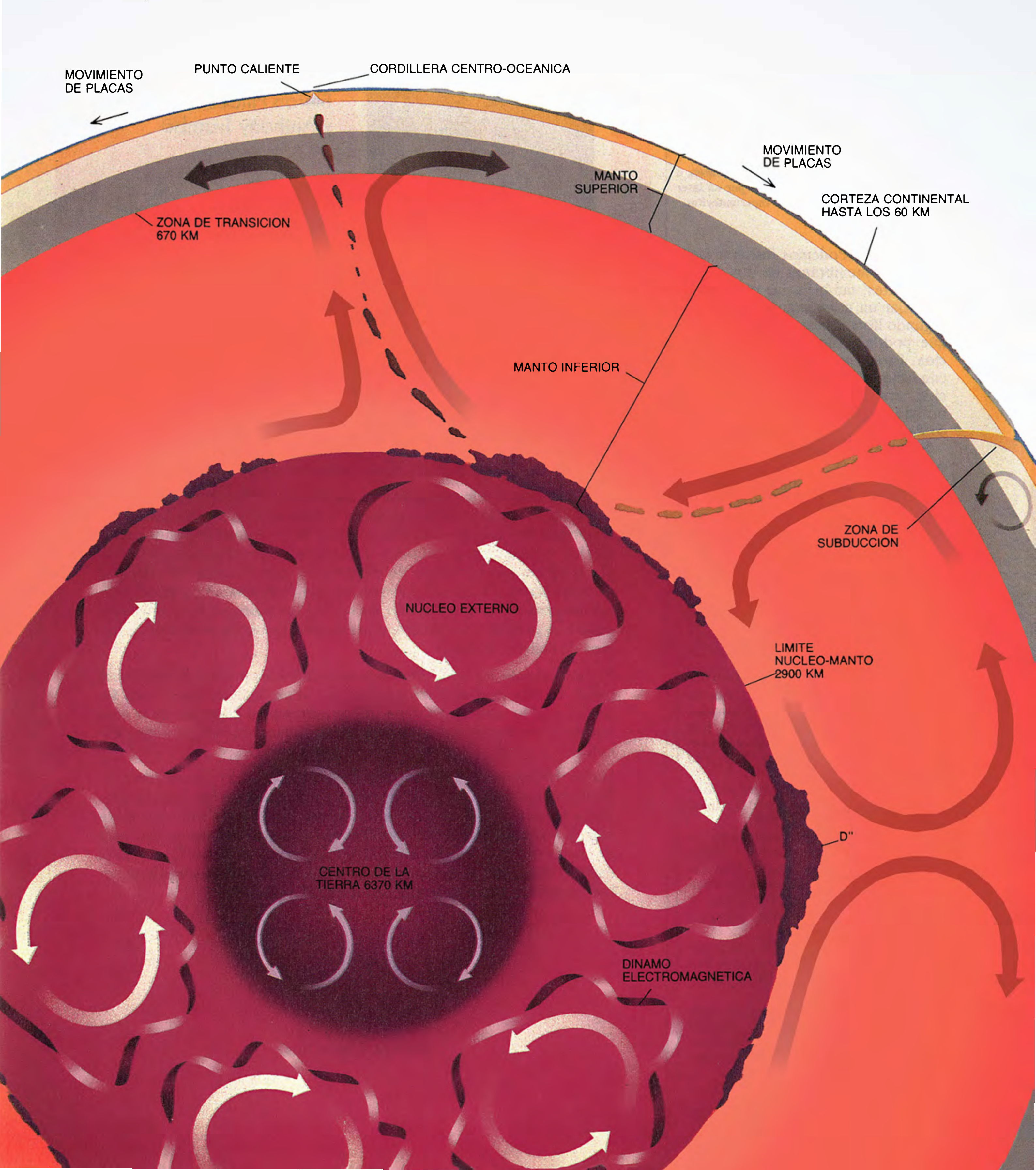
INTRODUCTION TO LASER PHYSICS AND LASER TISSUE INTERACTIONS. J. Stuart Nelson, William H. Wright, John Eugene y Michael W. Berns en *Endovascular Surgery*. Dirigido por Wesley S. Moore y Samuel S. Ahn. Harcourt Brace Jovanovich, 1989.

IEEE JOURNAL OF QUANTUM ELECTRONICS: SPECIAL ISSUE ON LASERS IN BIOLOGY AND MEDICINE, vol. 26, n.° 12; diciembre de 1990.

LASER MICROBEAM AS A TOOL IN CELL BIOLOGY. Michael W. Berns, William H. Wright y Rosemarie Weigand Steubing en *International Review of Cytology*, vol. 126, págs. 1-44; 1991.

MIRANDO HACIA DENTRO

Corey S. Powell



Los geofísicos disponen de técnicas nuevas que les permiten conocer las estructuras escondidas en las entrañas de la Tierra. Sus descubrimientos ofrecen incluso la posibilidad de mirar hacia el futuro del planeta.

OCEANO
LITOSFERA 100 KM
ASTENOSFERA 350 KM
CORDILLERA CENTRO-OCEANICA
MOVIMIENTO DE PLACAS
PUNTO CALIENTE

EL CALOR que escapa del interior de la Tierra la mantiene en constante agitación. Su campo magnético lo generan retorcidas corrientes de la parte externa de su núcleo de hierro fundido. Pausados y poderosos movimientos convectivos del manto silíceo desplazan los continentes y provocan casi toda la actividad geológica superficial. Los materiales pueden circular desde la parte inferior del manto (la región D") hasta su cima, por lo que, en algunas plumas de roca caliente arrastran pedazos de núcleo hasta la superficie (zona superior de la ilustración). *Cambios súbitos en la estructura y composición de los minerales a 670 kilómetros de profundidad podrían generar una convección en capas* (a la derecha). *Los números indican profundidades por debajo de la superficie terrestre.*

"He mirado a través de miles de kilómetros hacia el interior de la Tierra", dice Gary A. Glatzmaier, del Laboratorio Nacional de Los Alamos, con indiferencia, mientras un poderosísimo superordenador le permite simular las circunstancias de sus profundidades. Existe allí un mundo en continua agitación, como un caldero hirviente, que intenta desprenderse de su calor interno. Grandes extensiones de roca fría rompen en superficie y se hunden lentamente en el cálido interior, al tiempo que plumas de roca caliente y fundida se elevan y expanden como cúmulos nubosos al acercarse a la superficie.

Glatzmaier es un geofísico de vanguardia, de los que tratan de comprender la compleja dinámica interna de la Tierra. Llevan éstos su investigación más allá de la delgada piel externa (de no más de 60 kilómetros de espesor) donde se encuentran las montañas, las fallas en movimiento, los volcanes activos y demás accidentes del terreno. Su sondeo de las capas subyacentes, el núcleo y el manto, que constituyen más del 99 por ciento del planeta, está poniendo de manifiesto los procesos que, en última instancia, dan forma a la superficie.

Los estudios de simulación muestran las corrientes que existen en el manto sólido y flexible, que se extiende hasta los 2900 kilómetros. En el mundo de la realidad, estas corrientes proceden con suma lentitud: tardan cientos de millones de años en completar un ciclo, pero su fuerza es tal que mueven continentes y modifican océanos. Hay otra actividad más profunda en el núcleo de la Tierra, constituido por una aleación de hierro.

En su parte externa, que es líquida, se producen flujos tortuosos, cuya velocidad multiplica un millón de veces la celeridad de los desarrollados en el manto y que generan el campo magnético protector de la Tierra. Ha comenzado ya a levantarse el mapa de la circulación de ese océano metálico subterráneo. Las técnicas sísmicas, muy refinadas, nos facultan para asomarnos al núcleo sólido, cuyo límite exterior se encuentra más allá de las tres cuartas partes del radio terrestre, de 6370 kilómetros.

La nueva imagen que se está obteniendo de la estructura interna de nuestro planeta resulta tanto más notable por lo poco que se conocía hasta ahora. Su núcleo, que constituye más de la mitad del diámetro del globo entero, no fue descubierto hasta 1906 y el ritmo de los descubrimientos iniciales fue de una lentitud exasperante. Pasaron 65 años hasta que se pudo decir con certeza que el núcleo se divide en una parte externa, líquida, y otra interna, sólida. La teoría de la deriva continental, que enlaza íntimamente con el concepto de grandes movimientos de convección en el manto, no obtuvo una base firme hasta los años sesenta.

Que la imagen sea cada vez más detallada se debe al progreso experimentado en tres áreas generales: el creciente poder de los ordenadores; la obtención más sistemática de datos gravitacionales, sísmicos y magnéticos; y la disponibilidad de dispositivos capaces de reproducir las temperaturas y presiones existentes en el centro de la Tierra. Se descifran las ondas generadas por los terremotos y se analiza el campo magnético terrestre para cartografiar el núcleo y el manto. Con elaborados yunques de diamante se estudia el comportamiento de las rocas a las temperaturas y presiones infernales que imperan en el centro del planeta.

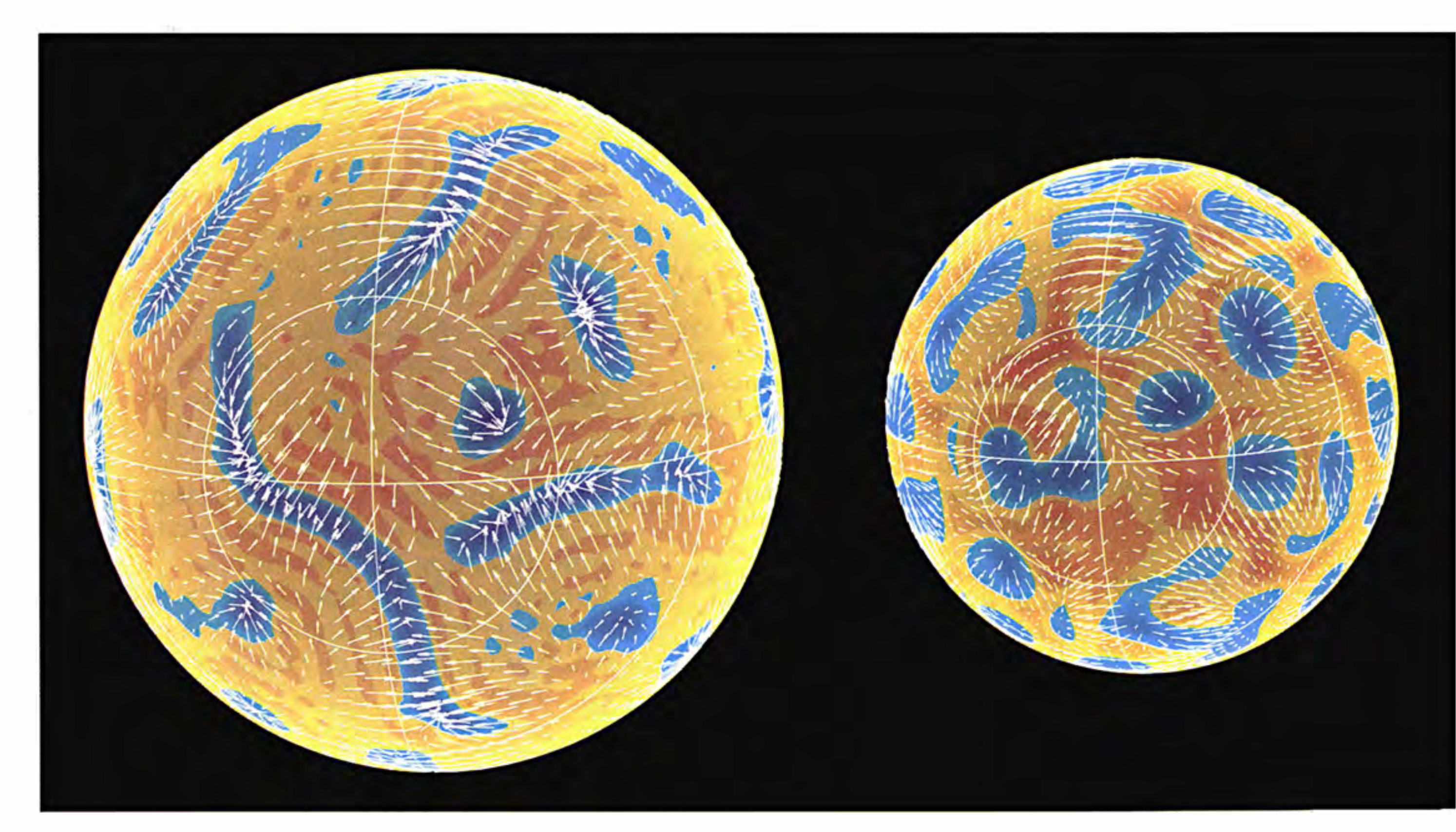
Y, conforme tan diversos elementos van encajando entre sí, se descubre que no sólo puede explicarse el comportamiento presente sino también formarse una idea de la evolución de la Tierra e incluso de su último destino, alejado miles de millones de años en el futuro. "Vivimos un momento muy excitante", dice Adam M. Dziewonski. En su despacho del Laboratorio Hoffman de la Universidad de Harvard, las paredes están cubiertas de mapas del interior de la Tierra y de las localizaciones de los terremotos, pero conserva un espacio libre para la pantalla de su proyector.

Las imágenes de las diapositivas que va desgranando muestran campos de formas amorfas, rojas y azules, que recuerdan los modelos creados por el equipo de Glatzmaier. Pero, en este caso, no se trata de simulaciones de ordenador sino de reconstrucciones del verdadero interior de la Tierra, producto de la tomografía sísmica, técnica innovadora que apareció a comienzos de los años ochenta.

Básase la tomografía en las ondas producidas por un terremoto, que se mueven a diferentes velocidades por diferentes partes de la Tierra; celeridad que viene condicionada por la densidad, composición, estructura mineral y grado de fusión. También influye la temperatura: las ondas viajan más deprisa en las regiones frías que en las calientes. La detección y el análisis de ondas sísmicas provenientes de todas las partes del globo permiten trazar un gran número de trayectorias por el interior de la Tierra.

Cuando esta información se somete a análisis por ordenador, pueden obtenerse imágenes tridimensionales de dicho interior. Los mapas de Dziewonski se adaptan al esquema de colores que se ha convertido en típico: el rojo denota ondas sísmicas bastante lentas, relacionadas con altas temperaturas, mientras que las regiones azules indican ondas rápidas, o bajas temperaturas.

A profundidades de unos 200 km, las zonas se presentan frías bajo los continentes y calientes bajo las cor-



1. CORRIENTES CALIENTES Y FRIAS ascienden y se hunden a través del manto. Las tres primeras imágenes son instantáneas de un modelo tridimensional de la Tierra convectiva. El color azul representa regiones más frías que la temperatura media del material; el naranja y el amarillo, más calientes que la media. Las flechas señalan la corriente superficial, a velocidades que no superan el par de centímetros al año. Los globos muestran cortes esféricos a través del modelo a profundidades de 430 km (*primer mapa de la izquierda*), 2020 km (*segundo por la izquierda*) y 2600 km, por

dilleras centro-oceánicas, donde las placas litosféricas se separan. Si se profundiza más, la estructura del manto no parece guardar ya relación con regiones específicas de actividad geológica, sino con movimientos a gran escala de los continentes. Hay materia caliente bajo Africa y en el centro del océano Pacífico. El anillo de roca fría que rodea el Pacífico subyace a todos los continentes, excepto Africa, como si hubiesen sido llevados a su situación actual por flujos que se hunden en el manto.

Los resultados de la tomografía parecen corroborar la noción intuitiva de que la convección del manto determina el movimiento de las placas tectónicas. Pero el estudio geofísico de las profundidades terrestres sigue ocasionando apasionados debates, que alcanzan su mayor intensidad entre quienes escudriñan los entresijos de la frontera entre el núcleo y el manto.

La Tierra dividida

El límite núcleo-manto representa la transición física y química más abrupta del interior de la Tierra. Allí es donde los silicatos del manto se enfrentan al núcleo de aleación de hierro. Las lecturas sísmicas indican que la roca de la base del primero es sólida y fluye con extremada lentitud, como si fuera vidrio. Por contra, la parte externa del núcleo tiene una consistencia cercana a la del agua. La diferencia de temperatura entre manto y núcleo puede ser de hasta 1000 grados Celsius.

Es más que probable que esta frontera guarde pistas valiosas sobre el funcionamiento y evolución de la Tierra, aunque encontrarlas trae de cabeza a los investigadores. La tomografía únicamente revela el modo en que las capas internas afectan el movimiento de las ondas sísmicas; la resolución de la mayoría de las técnicas disponibles es pobre en ese límite fronterizo, que abreviaremos por LNM. "Han estado circulando un montón de ideas disparatadas porque no podían verificarse", medita Don L. Anderson, catedrático de geofísica del Instituto de Tecnología de California y uno de los pioneros de la tomografía sísmica, aunque añade "... hasta ahora".

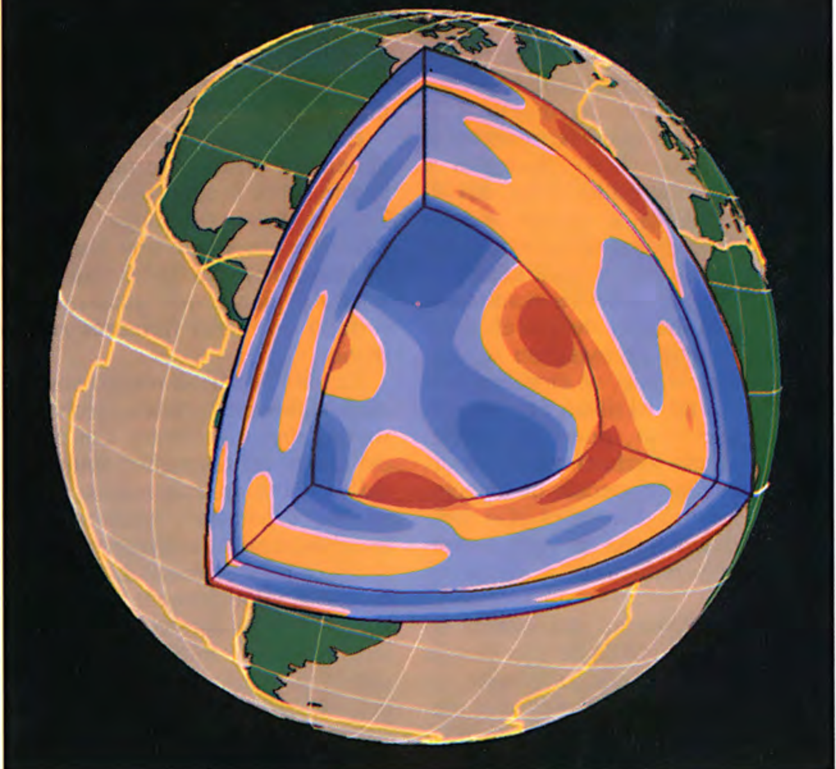
Cierto es que no puede llegarse hasta el núcleo, pero lo que sí puede hacerse ahora es traer el núcleo hasta el laboratorio, gracias a instrumentos de investigación de altas presiones y, muy en especial, a la célula de diamante en forma de yunque. En el Laboratorio Geofísico de la Institución Carnegie en Washington, en la Universidad de California en Berkeley y en algunos otros lugares pueden reproducirse las condiciones físicas del centro de la Tierra durante períodos dilatados, si bien a una escala muy pequeña (el tamaño de las muestras no supera una décima de milímetro).

Los estudios a altas presiones son cruciales para descifrar las complejas reacciones químicas que se producirán, a buen seguro, en el LNM, donde temperaturas entre los 3000 y los 4500 grados Celsius y presiones más de un millón de veces superiores a la del nivel del mar hacen que las sustancias interactúen de manera insólita y no bien comprendida.

Entre quienes intentan utilizar los indicios experimentales para ahondar en el conocimiento del LNM, uno de los más vivaces es Raymond Jeanloz, de Berkeley. En alusión a sus esfuerzos por deducir las reacciones químicas que ocurren en ese límite, David J. Stevenson, de Caltech, le califica de "explorador experimental". Jeanloz y Elise Knittle, ahora en la Universidad de California en Santa Cruz, han sugerido que esta zona es una barrera porosa que permite un grado de interacción química considerable entre ambas partes limítrofes.

Basan esta conclusión en un experimento que realizaron. Cargaron la cavidad del yunque de diamante con





encima mismo del límite núcleo-manto (*tercer mapa*). Las corrientes frías, descendentes, dominan el proceso convectivo; el material caliente desplazado se expande a medida que se va elevando. La tomografía proporciona una imagen tridimensional (*derecha*) de velocidades sísmicas desde la cima del manto hasta el límite núcleo-manto. Las regiones lentas (*rojo*) y rápidas (*azules*) vienen a corresponder, respectivamente, con material caliente, ascendente, y material frío, descendente. Los patrones de flujo del manto superior difieren de las pautas seguidas a profundidad mayor.

ONDA P
ONDA
DIRECCION DE PROPAGACION
ONDA S
ONDA
EPICENTRO
NUCLEO INTERNO
NUCLEO EXTERNO
MANTO
D
A
B
C
ZONA DE SOMBRA DE LA ONDA S

Las ondas sísmicas revelan el funcionamiento del interior de la Tierra

Las ondas sísmicas producidas por un terremoto o una gran explosión sirven de sonda de la estructura interna de la Tierra. Las ondas P son pulsos de compresión a través de la Tierra, análogas a las ondas sonoras del aire. Se propagan por el interior del globo. Las ondas S son deformaciones transversales de la Tierra sólida. Viajan a través de un material elástico y no pueden cruzar directamente el núcleo externo líquido. R. D. Oldham dedujo la existencia del núcleo en 1906, al observar la existencia de una zona de sombra para las ondas sísmicas S.

Las variaciones de la temperatura, presión y composición del interior pueden alterar la velocidad de las ondas y provocar la deformación o la reflexión de éstas. Las reflexiones se agudizan en la superficie, en el límite núcleo-manto (LNM) y en la superficie del núcleo interno sólido. El estudio de los tiempos de llegada de las ondas que han recorrido diferente camino a través de la Tierra nos faculta para inferir las propiedades físicas tridimensionales del interior de la Tierra.

Algunas ondas atraviesan el manto (*A*); otras, manto y núcleo externo (*B*), o también manto, núcleo externo y núcleo interno (*C*). Las ondas que se reflejan una o más veces en el interior de la Tierra aportan información muy valiosa para conocer su interior. En una técnica que se emplea para examinar estructuras del fondo del manto, se sustraen los tiempos de llegada de las ondas S que se reflejan en el LNM (*D*) de los tiempos de las ondas S que llegan al mismo sitio directamente. La comparación de las ondas P que se reflejan en la superficie del núcleo con las que lo atraviesan corrobora la variación de la altitud del LNM de un lugar a otro.

trocitos de minerales de sílice (la mejor suposición que pudieron hacer sobre la composición del manto inferior) y de hierro (para representar el núcleo externo). Los calentaron a temperaturas parecidas a las del límite entre ambos, al tiempo que los sometían a presiones de hasta 800.000 atmósferas, tras de lo cual examinaron la pequeña muestra resultante. Encontraron indicios de que el silicio y el hierro, que permanecen separados en la superficie, se entremezclan en este entorno. El silicio y el oxígeno podrían encontrar, pues, su camino desde el manto hacia el núcleo externo mediante la formación de aleaciones con el hierro, las cuales irían cambiando gradualmente la composición del núcleo y reducirían su densidad media.

Como la densidad del núcleo es bastante menor que la de una mezcla pura de hierro y níquel, es opinión común que también debe andar por medio algún otro componente más ligero. Muchos piensan que la existencia de impurezas se remonta a los tiempos de formación de la Tierra, pero Jeanloz ha introducido la variante de que la composición del núcleo esté sometida a continua evolución conforme se disuelven en él pedazos del manto silíceo.

En condiciones adecuadas, estos materiales podrían formar una capa aparte sobre la superficie del núcleo. Thorne Lay, sismólogo de Santa Cruz, ha observado una disminución de las velocidades de las ondas en tal región y piensa que pudiera ser resultado de restos flotantes ricos en silicio.

La idea de que el núcleo conste de capas diferentes ayuda a resolver otro rompecabezas geofísico. Los estudios del campo geomagnético apuntan a la existencia de un flujo regular en la superficie del núcleo, mientras que los estudios sísmicos realizados por Dziewonski y otros sugieren irregularidades de varios kilómetros de altura. Quizá haya una capa accidentada, hecha de una mezcla de núcleo y manto, que cabalgue sobre la superficie alisada del primero.

Pero pudiera ser que la estratificación del núcleo fuese una ilusión. Jeremy Bloxham, de Harvard también, se muestra escéptico sobre una explicación tan "ad hoc". Lleva años correlacionando datos históricos del campo magnético y cree que "resulta cada vez más difícil" reconciliar sus resultados con la existencia de una capa superior diferenciada del núcleo.

Y la cosa se complica aún más a este lado del límite núcleo-manto. Hay indicios cada vez más claros de la presencia de estructuras misteriosas por encima mismo del LNM, lo que se conoce como región D″. Diversas investigaciones, entre ellas las de Lay, señalan que la velocidad de las ondas sísmicas cambia a unos 200-300 kilómetros por encima de la divisoria LNM, lo que supondría una capa de transición, mientras que otras muestran una delimitación casi nítida. Por ello, aboga Lay, el límite núcleo-manto tiene una estructura muy variable, aparente en algunos sitios y no en otros.

Si se intenta entrar en detalles, la armonía se convierte en un conjunto casi ilimitado de teorías dispares. Jeanloz sostiene que todo lo que desciende debe subir: al igual que el manto puede disolverse en el núcleo, material de éste puede escapar hacia aquél. Apunta la idea de que el hierro del núcleo podría ser arrastrado por

acción capilar hacia el manto, donde formaría agregados de aleación de hierro o rocas silíceas ferruginosas. Variaciones locales de la temperatura o la composición química explicarían la granularidad regional de la zona D″.

Stevenson no discute que la D″ pueda recibir material del núcleo, pero señala otras posibles fuentes de mezcla. En el fondo del manto "también podría haber una especie de basurero", al que, por ejemplo, fuesen a parar los restos de viejas placas litosféricas subducidas. O podría ser el destino final de material primigenio suficientemente denso para hundirse en el manto pero demasiado ligero para incorporarse al núcleo. Muchos de sus colegas se hacen eco de esta idea.

Lay añade que es muy probable que la temperatura y la presión influyan decisivamente en la estructura física de la D″. Al ser tan grande el gradiente de temperatura de la demarcación o LNM, el calor proveniente del núcleo podría provocar inestabilidades periódicas en la D″, cuyos materiales calientes y más ligeros podrían formar una pluma. Peter Olson, de la Universidad Johns Hopkins, y otros han desarrollado modelos que demuestran que las plumas del manto tendrían la forma de estrechos conos de material caliente que, al alcanzar la superficie, crearían puntos calientes volcánicos.

Son éstos regiones de actividad volcánica que permanecen fijas durante millones de años, mientras las placas litosféricas que soportan los continentes se deslizan sobre ellos. Un punto caliente bajo el océano Pacífico creó la nítida línea de las islas Hawai; su presencia se hace sentir todavía en las frecuentes erupciones del volcán Kilauea. No cabe duda de que los puntos tienen su origen en una fuente estacionaria y, por tanto, situada a enorme profundidad, que descarga a través del suelo marino o de los continentes según derivan lentamente por el manto.

¿Uno o dos mantos?

Si las plumas se originan en el LNM, deberían transportar trazas de material del núcleo hasta la superficie. Los gases expulsados a la superficie por los volcanes contienen helio 3, una reliquia procedente de la formación de la Tierra, que podría haber estado atrapada durante miles de millones de años en el manto, aunque Stevenson piensa que, en parte, también podría haber pasado desde el núcleo hacia la región D″, para ser luego arrastrada hacia arriba. "No hay manera de probarlo," admite, "pero parece razonable pensar que algunos átomos de la superficie de la Tierra hayan estado en el núcleo en un tiempo geológicamente reciente."

Anderson, por su lado, tiene una perspectiva muy diferente sobre el destino de los materiales de la D″. Considera que el material del fondo del manto es demasiado denso para ascender hasta la corteza. Podrían formarse pequeñas células de convección que, si la capa se calentase lo suficiente, ascenderían para volver a hundirse en el manto inferior, lo que, según dice, "es una historia bastante diferente de la de plumas que ascienden hasta la superficie".

El curso que siguen las plumas a través de la Tierra depende, por supuesto, de la circulación del manto; lo que, a su vez, es cuestión muy disputada. Hay quienes mantienen que el manto entero participa en el proceso de mezcla, gracias a unas enormes corrientes de convección que abarcan desde donde limita con el núcleo hasta la base de la corteza. Otro grupo defiende un manto compuesto por dos capas al menos, dotadas de convección independiente. La estimación global de Stevenson es que un 80 o 90 por ciento de los geofísicos se inclinan por la idea de una convección profunda en todo el manto. Sin embargo, Russell J. Hemley, físico de altas presiones del Laboratorio Geofísico, cree que la tendencia actual va en el otro sentido, apoyada en un número creciente de datos experimentales que abogan por un cambio apreciable, en la composición del manto, de acuerdo con la profundidad.

Son muchos quienes citan la distribución de puntos calientes como uno de los argumentos que sustentan la convección profunda. Dziewonski, en colaboración con Mark A. Richards, de Berkeley, y Bradford H. Hager, del Instituto de Tecnología de Massachusetts, ha puesto de manifiesto que los puntos calientes de la superficie tienden a situarse sobre regiones del LNM a las que los mapas de tomografía sísmica señalan como bastante calientes. ¿Podemos inferir que los puntos calientes dan salida a estos excesos de temperatura y que existe una conexión directa entre la demarcación y los volcanes de la superficie?

Convección

Las simulaciones por ordenador desarrolladas por Glatzmaier y otros pioneros, así David Bercovici, de la Universidad de Hawai, y Gerald Schubert, de la de California en Los Angeles, dan a entender que las fuentes de calor interno (desintegración radiactiva) producen convección a gran escala, mientras que el calentamiento desde abajo crearía plumas. Basándose en esas simulaciones, Geoffrey Davies, experto en puntos calientes de la Universidad Nacional de Australia, propone que la conveccion del manto disipa su propio calor y que las plumas dan alguna salida al calor del núcleo. Si éstas no provinieran de tal interfase, argumenta, el calor que surge del núcleo generaría fuertes corrientes ascendentes, fenómeno que no se ha observado.

Thomas Jordan, del MIT, también cree tener pruebas de que la convección se extiende de forma continua por todo el manto. No utiliza técnicas tomográficas sino que centra su atención en el estudio de las ondas sísmicas reflejadas por el LNM. Su propuesta es la de un "sistema de cinta transportadora": enormes células de convección desplazan los continentes en la superficie al tiempo que, en una especie de extraño mundo paralelo, agitan las masas irregulares de material situadas por encima del límite núcleo-manto.

Quienes no creen en el modelo de circulación profunda ven las cosas de modo algo diferente. Los experimentos de Jeanloz y Knittle implican que el manto inferior es demasiado denso para tener la misma composición que el manto superior. La inyección de algo de hierro del núcleo al manto inferior incrementaría la densidad hasta los niveles observados. La conclusión de Jeanloz es que no puede haber convección profunda, porque impediría esta segregación química, aparte de que el material denso del manto inferior tampoco podría elevarse mucho.

Pero ésa es una línea de argumentación que no convence a todo el mundo. Algunos objetan que la perovskita, el mineral predominante en el manto inferior, es bastante más denso a altas presiones de lo que conceden los experimentos de Jeanloz. Apoyado en sus propias pruebas con él en células de diamante, Hemley sostiene que hay que modificar los resultados de Jeanloz en lo que a la densidad se refiere.

El comportamiento del manto parece volverse especialmente complicado en la zona de transición, entre los 400 y los 670 kilómetros de profundidad, donde cambia bruscamente la velocidad de las ondas sísmicas. Su extensión parece venir determinada por las presiones a las que los minerales del manto superior (en especial el olivino) se deforman hasta adquirir nuevos estados cuasiestables (perovskita y wüstita). Las placas litosféricas absorbidas por el manto dan la sensación de curvarse y desviarse cuando

alcanzan la zona de transición, como si hubiesen tropezado con una frontera impenetrable.

Los partidarios de la convección estratificada sostienen que un bucle de convección del manto se mueve entre el LNM y la parte inferior de la zona de transición de 670 km, mientras que otro grupo lo hace entre la zona de transición y la periferia del manto. Según este punto de vista, los materiales situados por debajo de los 670 kilómetros no alcanzarían la superficie y lo que pueda suceder en el límite núclo-manto sólo afectaría indirectamente al manto superior.

Anderson considera que "no existen pruebas" de que ningún material que se halle a más de 670 km tenga nada que ver directamente con los fenómenos de la superficie. Sería mucho más sencillo, agrega, que las plumas se formasen a 670 km, gracias al flujo de calor de la mayor parte del manto, que en el LNM, con el núcleo como única fuente de calor. Para él, la composición de las lavas de diferentes volcanes denuncia una mezcla incompleta del manto, por lo que no es necesario recurrir a material arrancado del núcleo.

Por el camino de enmedio

De todos modos, la geofísica de las profundidades de la Tierra sigue siendo un dominio poco asentado y lo que a primera vista parecen enconados conflictos no es raro que se desvanezcan en tolerables diferencias de interpretación. Lay admite que las regiones calientes del fondo del manto podrían actuar como las placas de una cocina y calentar las zonas superiores correspondientes, aunque no hubiera mezcla entre ellas. Jeanloz no bromea cuando dice que "los modelos de convección global del manto podrían ser correctos", aunque lo más probable es que "la realidad ande por algún punto intermedio". Por poner un ejemplo, el manto podría estar estratificado químicamente de modo tal que permitiese que determinados materiales ascendieran por encima o se hundieran por debajo del límite de 670 km.

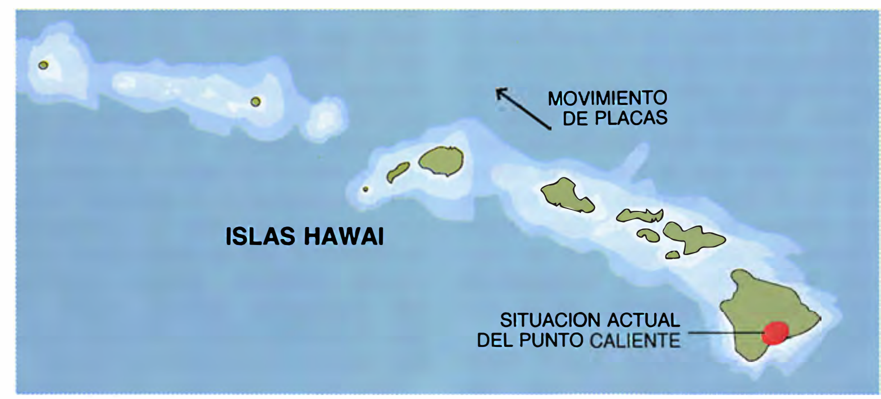
Philippe Machetel y Patrice Weber, del Grupo de Investigación de Geodesia Espacial en Toulouse, han desarrollado un detallado modelo de ordenador para explorar esta posibilidad. No es como el modelo tridimensional de la Tierra creado por Glatzmaier, Schubert y Bercovici, ya que el grupo francés sólo puede ver un corte bidimensional, pero así dejan libre una capacidad de cómputo que les permite explorar mayor cantidad de propiedades físicas del manto real, como las relacionadas con cambios de temperatura o presión. Su modelo utiliza ecuaciones verosímiles del comportamiento de los minerales en la zona de transición y permite observar cómo van acumulándose materiales a 670 km bajo la superficie hasta que, cuando se alcanza determinado estado crítico, atraviesan la barrera en uno u otro sentido.

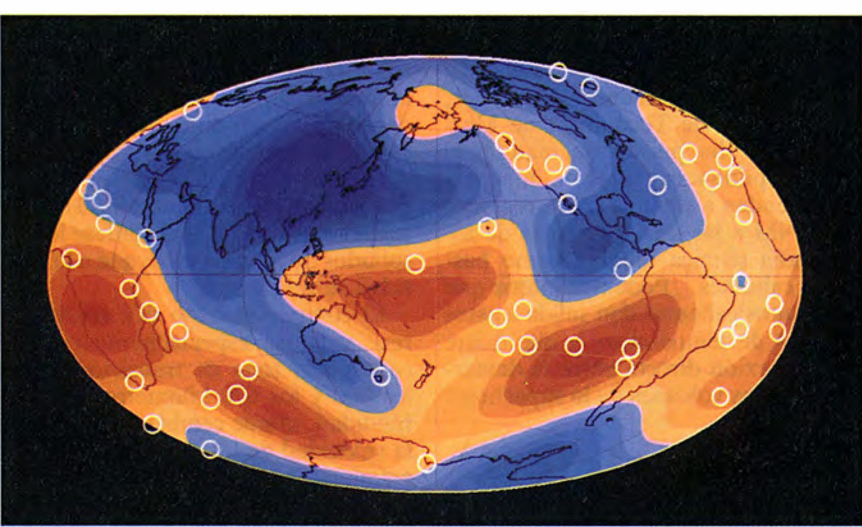
La forma en que se efectúe la convección tiene grandes implicaciones en la evolución química de la Tierra. Para Hager, el manto profundo es muy viscoso, por lo que cualquier mezcla completa debe producirse muy despacio. A tenor del ritmo de flujo de calor que atraviesa el manto, del arrastre sobre las placas litosféricas y de otros datos semejantes, Hager y sus colegas han calculado que el manto inferior es aproximadamente 30 veces más viscoso que el superior.

Si están en lo cierto, el manto profundo debe circular con la suficiente lentitud (quizás una vuelta cada mil millones de años) para que grandes masas permanezcan casi inalteradas y retengan todavía, en buena medida, la misma composición que tenían cuando la Tierra se formó, hace 4500 millones de años. Esas bolsas podrían constituir el "depósito primitivo" que necesitaba Stevenson para explicar las bocanadas de helio 3 que emiten los volcanes. Anderson abunda en ello cuando sostiene que fuentes del manto darían cuenta fácilmente de los datos observados. Pero a Stevenson le preocupa que la convección fuera más vigorosa en el pasado, cuando la Tierra estaba más caliente, el manto era menos viscoso y mayor el escape de calor.

En las entrañas del núcleo

Los interminables esfuerzos de la Tierra por deshacerse de su calor interno afectan tanto a su parte más recóndita —el núcleo de aleación de hierro— como a la superficial que acabamos de considerar. En esta última, la circulación viene condicionada por las regiones frías y calientes del manto inferior circundante, rela-





2. UN PUNTO CALIENTE DEL MANTO atravesó la corteza terrestre, creando la cadena lineal de las islas Hawai (*arriba*). Los puntos calientes permanecen estacionarios, mientras que las placas litosféricas que los recubren se deslizan poco a poco. El mapa de velocidades sísmicas de la base del manto (*abajo*) manifiesta que casi todos los puntos calientes (*círculos blancos*) se asientan sobre regiones a altas temperaturas del manto inferior (*color rojo*), prueba de que los puntos calientes podrían originarse desde una capa situada encima mismo del núcleo caliente de la Tierra.

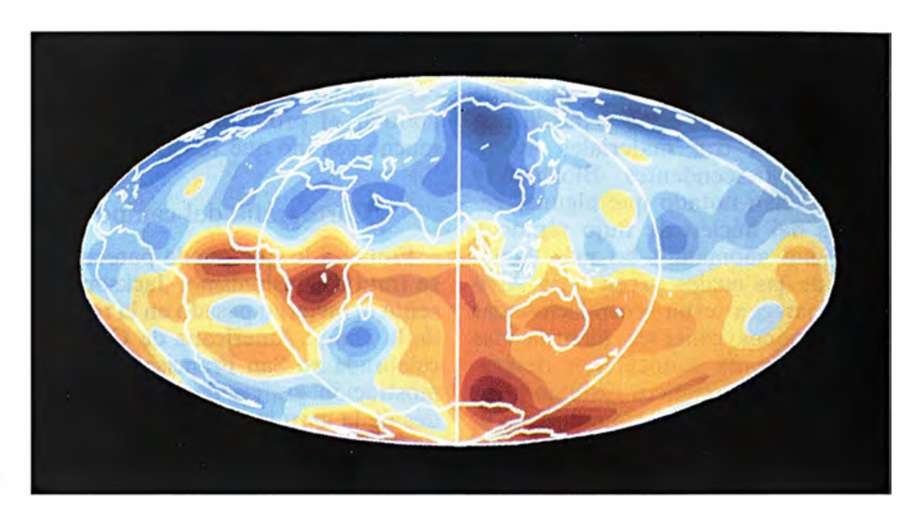
tivamente estacionario si se tiene en cuenta que la velocidad de flujo del núcleo externo es de varios kilómetros al año, o sea, como un millón de veces más rápida. Por su base, el núcleo externo absorbe calor del enigmático núcleo interno.

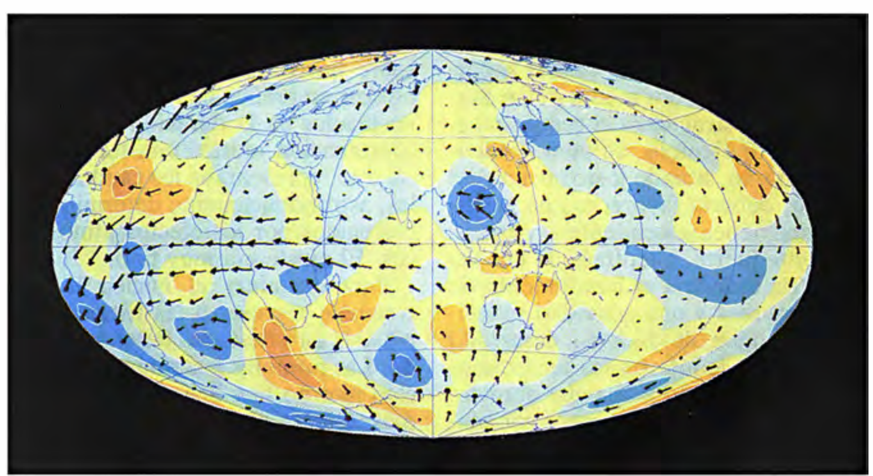
A medida que el núcleo externo se va enfriando lentamente, los cristales de hierro se separan del líquido fundido y se depositan en la parte superior del núcleo interno, con liberación de calor en el proceso. Es así como el núcleo interno ha estado creciendo sin cesar a lo largo de la historia de la Tierra. No obstante su gran temperatura, casi la misma que la de la superficie del Sol, las enormes presiones que existen en el centro del planeta mantienen sólido el núcleo interno. Pero incluso aquí el calor se va difundiendo hacia afuera. Estudios tomográficos recientes confirman que el núcleo interno, al igual que el manto, presenta un ritmo de convección lento a medida que se enfría.

La rapidez del flujo del núcleo externo puede percibirse con claridad lejos de la Tierra. Hay amplio acuerdo en que es él quien genera una corriente eléctrica y alimenta su campo magnético. Siendo así, debería poder "escucharse" el "chapoteo" del núcleo externo sólo con mirar la orientación de una brújula muy sensible. En principio, debería ser también posible trabajar en sentido inverso y reconstruir las condiciones que allí se dan midiendo en superficie la estructura y el comportamiento del campo magnético terrestre.

Ni más ni menos que esto se propuso hacer Bloxham, de Harvard. El mundo que él observa es bastante más activo que el de los sismólogos, ya que los cambios del núcleo externo se producen en decenios y no en decenas de millones de años, como los del manto. Durante los últimos siglos, la componente norte-sur (dipolo) del campo magnético terrestre se ha debilitado de forma apreciable y el campo, en su conjunto, ha derivado hacia el oeste. Si se consideran períodos de tiempo más largos, se producen disminuciones y recuperaciones periódicas del campo, intercambiándose los polos norte y sur.

Bloxham compara modestamente su trabajo con la predicción del tiempo. Cuando lo observamos en cierto instante obtenemos sólo datos inmediatos sobre la temperatura, la humedad o el viento, pero si lo hacemos en el transcurso de unos días dispondremos de valiosa información sobre el comportamiento de la atmósfera, el flujo de los vientos, el movimiento de las tormentas, etc. De la misma manera, registrando los cambios graduales del campo magnético terrestre es posible inferir el clima del núcleo externo.





3. NUCLEO EXTERNO DE HIERRO FUNDIDO, sujeto a una circulación vigorosa que generará el campo geomagnético. Arriba se ilustra el campo magnético de la superficie del núcleo. El azul señala que el flujo magnético cursa hacia el norte; el color rojo y el naranja, el curso meridional de la circulación. El mapa inferior recoge los movimientos que, en el núcleo, dan origen a la distribución irregular del campo. Las flechas muestran la dirección y la velocidad: los flujos más rápidos avanzan 30 km al año. El color naranja denuncia un proceso de afloramiento; el azul, de subsidencia.

Para abordar el problema, Bloxham formó equipo con otros investigadores del campo magnético: David Gubbins, de la Universidd de Leeds, y Andrew Jackson, de la Universidad de Oxford. Comenzaron por estudiar larga y detenidamente registros de medidas del campo magnético desde finales del siglo XVII. Después utilizaron diversas técnicas matemáticas para proyectarlas sobre la superficie del núcleo. Obtuvieron así mapas de flujo magnético en la frontera núcleo-manto que abarcan casi tres siglos. Un segundo turno de cálculos les permitió reconstruir el flujo en la superficie del núcleo que produce el campo magnético observado.

Antaño los libros de texto solían representar el campo magnético terrestre mediante un imán ingente en forma de barra, situado en el interior de la Tierra y ligeramente desplazado del centro. Los mapas de Bloxham son, por contra, gráficos coloreados de líneas de contorno, con los rojos y azules de los geofísicos, que definen regiones de polaridad positiva y negativa. Hasta la división entre polo norte y polo sur ha perdido su encantadora sencillez.

Regiones de polaridad inversa invaden ambos hemisferios. La situación asimétrica del campo global parece provenir de la desigual distribución de las manchas del núcleo, regiones de polaridad anómala que cambian mucho en unas decenas de años como reflejo del rápido ritmo de circulación del núcleo externo.

Pero, según Bloxham, "la erupción de las manchas del núcleo está casi con seguridad influenciada por la situación de los puntos calientes del manto". Su movimiento es tan lento que, desde la perspectiva del núcleo, el manto está casi inmóvil, pero las regiones calientes podrían promover una convección de notable vigor en el material subyacente del núcleo.

Los mapas de tomografía sísmica muestran regiones calientes y frías del manto, situadas por encima mismo de la superficie del núcleo y puede que relacionadas con materiales ascendentes y descendentes. Bloxham y Jackson han notado que algunas regiones del núcleo de flujo rápido y polaridad anómala —las más llamativas de las cuales yacen bajo Africa— parecen estar correlacionadas con regiones calientes identificadas por tomografía sísmica. Por desgracia, los mapas sísmicos resultan poco fiables a tan grandes profundidades y, en consecuencia, cualquier inferencia sobre la temperatura debe tomarse con extrema precaución.

Bloxham vio que, si invertía el problema, podía producir sus propias medidas de temperatura en la superficie del núcleo, por medio de un examen de la estructura del campo magnético. Por ejemplo, un punto caliente del manto, situado justo encima del núcleo, debería favorecer una actividad magnética ascendente y particularmente fuerte, mientras que una región fría induciría el hundimiento de las corrientes y una actividad reducida. En unión de Jackson desarrolló esta idea y dedujeron las temperaturas en el LNM. El mapa que obtuvieron coincide a grandes rasgos con mapas de temperatura levantados con tomografía sísmica, aunque quedan discrepancias suficientes para no hacer olvidar que ambas formas de abordar el problema tienen todavía mucho camino que recorrer.

Cartografía del campo

Frente a lo que pueda parecer, no se trata de ilusionismo. Jackson presentó el invierno pasado en la reunión de la Unión Americana de Geofísica celebrada en San Francisco, una demostración gráfica de las posibilidades de la cartografía del campo magnético. Si se produce éste por corrientes en el núcleo, las corrientes deben rozar contra la frontera interior del manto y afectar ligeramente la rotación de la Tierra. Pasó rápidamente sobre un conjunto de cálculos, para ofrecer un gráfico de los cambios pronosticados en la duración del día. A continuación mostró una transparencia que representaba los cambios reales, astronómicamente determinados, corregidos por los efectos atmosféricos. El ajuste era muy fino.

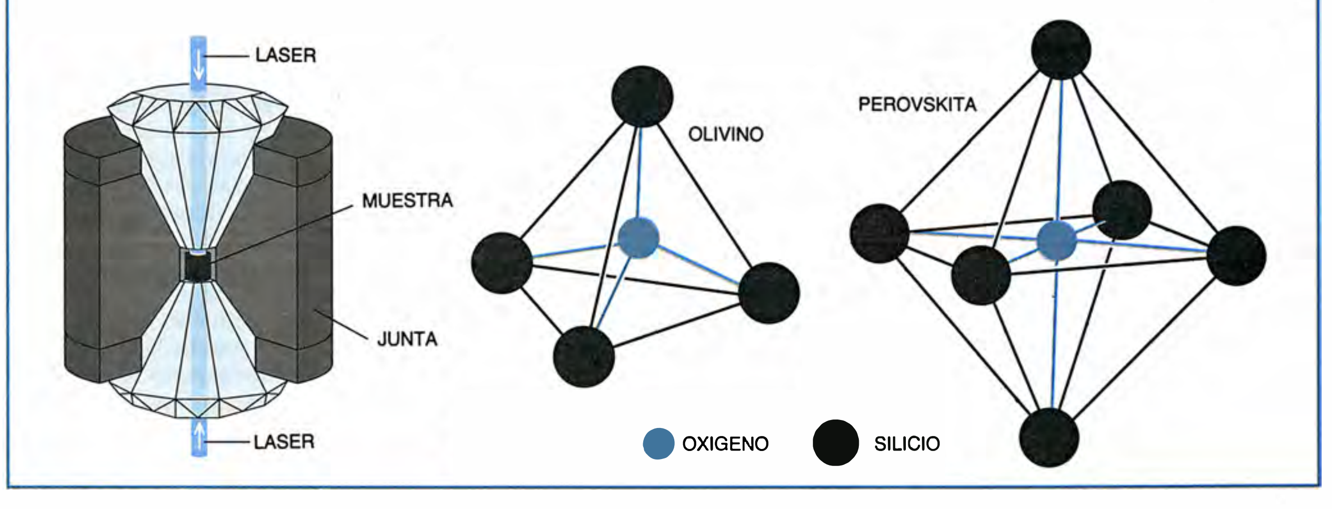
Para examinar el núcleo interno conviene regresar a las técnicas empleadas en el estudio del manto sólido. Las primeras instantáneas del núcleo interno han llegado de los estudios tomográficos centrados en las ondas sísmicas que implican vibraciones de la Tierra entera. Los terremotos o las grandes explosiones provocan que el globo vibre como una campana. Si la Tierra fuera uniforme en todas las direcciones, todas las partes vibrarían al unísono. Pero es un tanto elíptica, y el manto varía de composición y de temperatura de unos lugares a otros, lo que complica la pauta de oscilación.

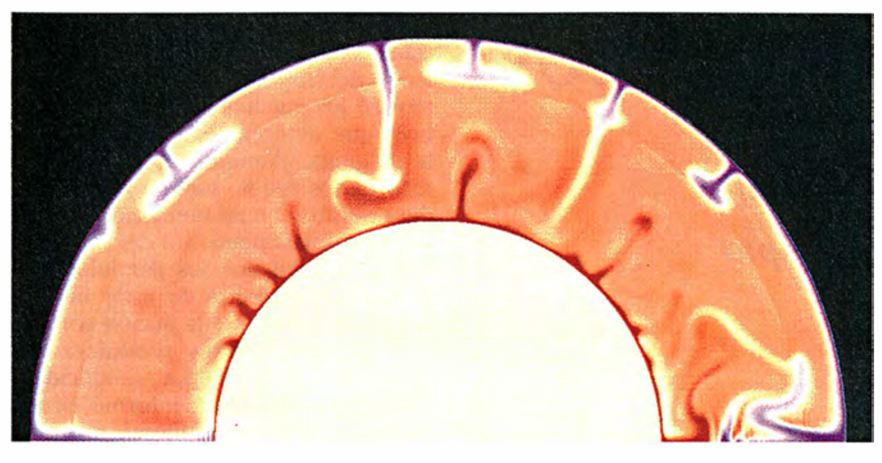
Atendidos esos factores, queda todavía sin explicar alguna distorsión en la vibración de la Tierra. A partir de los patrones de oscilación detectados, el sismólogo John H. Woodhouse, de Oxford, llega a la conclusión de que esa distorsión se origina en el núcleo interno. (Sismólogos del Caltech ponen en duda que eso sea cierto.) Woodhouse aduce una "explicación obvia" para la distorsión del núcleo. Ya que los cristales de hierro son en sí mismos asimétricos, las ondas sísmicas viajan más deprisa en el sentido de alineamiento de los cristales que en sentido perpendicular. Un alineamiento preferente de los cristales de hierro en el núcleo interno podría explicar la señal observada.

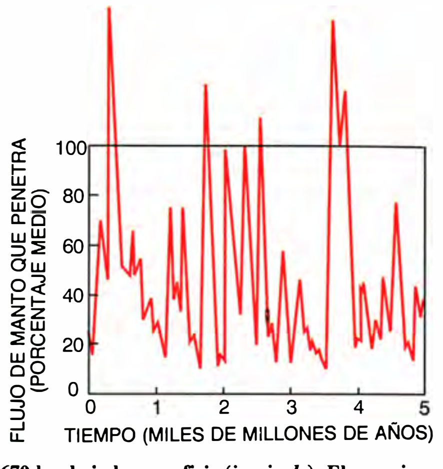
Pero, ¿en virtud de qué se alinean los cristales? Para Jeanloz se trataría de un efecto de la convección en el núcleo interno. Los movimientos serán, admite, lentísimos, de escasos centímetros al año. Lo que no cons-

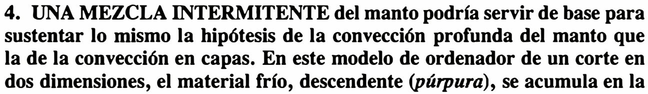
La geofísica trabaja bajo presión

Los dispositivos de altas presiones permiten la simulación en el laboratorio de las condiciones que existen en el manto e, incluso, las del núcleo. La cámara de diamante en forma de yunque (*izquierda*) incorpora dos diamantes, entre cuyas caras paralelas hay sólo una décima de milímetro. Los diamantes pueden presionar una muestra hasta cuatro millones de veces la presión atmosférica. Gracias a la transparencia del diamante, podemos observar la muestra mientras está bajo presión. Los diamantes se prestan también al calentamiento de la muestra por láser y a la investigación de su estructura con el bombardeo de haces intensos de rayos X. Se trata de cotejar los experimentos de laboratorio con las propiedades sísmicas observadas en las rocas del manto. El manto superior está constituido, sobre todo, por minerales con estructura de olivino, en que cuatro átomos de oxígeno rodean a un átomo de silicio (*centro*). Los experimentos a altas presiones han revelado que el olivino se reordena en perovskita, de estructura más densa (*derecha*), desde los 670 km de profundidad. La mayor parte del volumen de la Tierra está formado por perovskitas.









4. UNA MEZCLA INTERMITENTE del manto podría servir de base para sustentar lo mismo la hipótesis de la convección profunda del manto que la de la convección en capas. En este modelo de ordenador de un corte en dos dimensiones, el material frío, descendente (*púrpura*), se acumula en la transición de mineral a 670 km bajo la superficie (*izquierda*). El empuje a su regreso de las corrientes ascendentes puede provocar que el material frío atraviese la zona de transición y se hunda. El porcentaje del manto que atraviesa la zona de transición oscila de forma caótica.

tituiría óbice para que tales corrientes justificaran la orientación verosímil de los cristales de hierro en el núcleo interno. Woodhouse concede que ese proceso de convección parece razonable. La convección alteraría algo la forma del núcleo interno (y, con ello, las corrientes y los cristales), lo que provocaría su alineación con el eje de rotación. La orientación del núcleo interno deducida de las oscilaciones sísmicas se encuentra en un margen de 20 grados respecto del eje de rotación de la Tierra.

La Tierra que fue y la que será

Desde el núcleo interno hasta la superficie, la actividad dinámica está determinada por el calor. Máquina térmica a la postre, la Tierra se irá deteniendo gradualmente.

La predicción del futuro es un asunto delicado. En particular, la distinción entre la convección profunda del manto y la convección estratificada tiene profundas implicaciones para el futuro de la Tierra. Un manto superior que no se mezcle con el inferior mantendría, cual manta aislante, el calor del globo. Una convección profunda aportaría material caliente hasta la base de la corteza, permitiendo que el manto se enfriara mucho más deprisa. En ese marco, la Tierra habría sufrido una gran transformación a lo largo de 4500 millones de años de historia.

Las huellas de su comportamiento en el pasado nos ofrecen pistas para explicar su futuro, siempre que se sepan interpretar. Norman Sleep, geofísico de Stanford partidario de la convección profunda, habla de una Tierra antigua donde el manto era más caliente, menos viscoso y de circulación más vigorosa. Si anduviera en lo cierto, el salto de temperatura en el LNM habría sido mucho menor y no habría habido plumas o puntos calientes, o hubiesen sido débiles. Sleep admite, empero, la falta de pruebas, anteriores a los últimos 1300 millones de años, relativas a la actividad volcánica debida a puntos calientes.

Anderson, partiendo de la misma consideración, llega a una conclusión muy distinta. Es muy difícil, arguye, identificar puntos calientes e inundaciones de basalto con una antigüedad superior a mil millones de años. Dentro de esta limitación temporal, no encuentra ninguna señal de caída de actividad de los puntos calientes, enfoque coherente con el modelo de convección estratificada.

En cuanto al futuro, Jeanloz opina que la Tierra persistirá otros 5000 millones de años. Anderson asiente, habida cuenta de que la Tierra se ha enfriado sólo 200 grados Celsius a lo largo de su vida, lo que atribuye en parte a la convección en capas. Sleep adopta una visión más pesimista; conforme se enfría el manto, aumenta en rigidez y pierde movilidad; y en los próximos 1000 millones de años, el movimiento de las placas disminuirá o cesará, predice. Los continentes se establecerán en posiciones fijas, y la erosión nivelará las grandes cadenas montañosas. Las plumas continuarán avanzando hasta la superficie y se convertirán en el principal medio por el que la Tierra libera su calor interno.

Es de presumir que no necesitemos miles de millones de años para comprobar o desmentir tales predicciones. Podríamos quizá contemplar ahora mismo el futuro de la Tierra... en Venus. La sonda *Magallanes* nos ha revelado un planeta salpicado de volcanes y desprovisto de signos alusivos a la tectónica de placas. Venus, como la Tierra futura, "parece estar dominado por grandes y lentas plumas," dice Sleep. Pero Venus debió de enfriarse más deprisa que la Tierra: posee una superficie muy caliente (por encima de los 450 grados Celsius), fenómeno que, en opinión de Sleep, mantuvo las rocas próximas a la superficie blandas y flexibles y facilitó, con ello, una circulación más completa de las partes externas del planeta. Venus pudo deshacerse así de su calor interno con mayor rapidez que la Tierra.

La tomografía sísmica, que ha proporcionado imágenes de la estructura de la tierra con grandes posibilidades, tiene sus críticos. Hay quien ve contradicciones en los mapas o discute el valor de los resultados de los modelos bidimensionales. De cualquier modo, una serie de eslabones liga la deriva de los continentes, los terremotos, el campo geomagnético y los volcanes. Sólo falta conocerla bien.

BIBLIOGRAFIA COMPLEMENTARIA

INSIDE THE EARTH: EVIDENCE FROM EARTHQUAKES. Bruce A. Bolt. W. H. Freeman, 1982.

LA TIERRA. Prensa Científica, S.A. Barcelona, 1987.

GLOBAL IMAGES OF THE EARTH'S INTERIOR. Adam M. Dziewonski y John H. Woodhouse en *Science*, vol. 236, págs. 37-48; 3 de abril de 1987.

STUDIES OF THE EARTH'S DEEP INTERIOR: GOALS AND TRENDS. Thorne Lay, Thomas J. Ahrens, Peter Olson, Joseph Smyth y David Loper en *Physics Today*, vol. 63, n.° 10, págs. 44-52; octubre de 1990.

EARTH'S CORE-MANTLE BOUNDARY: RESULTS OF EXPERIMENTS AT HIGH PRESSURES AND TEMPERATURES. Elise Knittle y Raymond Jeanloz en *Science*, vol. 251, págs. 1438-1443; 22 de marzo de 1991.

Juegos matemáticos

Un viaje de Swift por terreno abrupto

Ian Stewart

En el verano de 1990, la directora de conservación bibliográfica de la Biblioteca Bodleyana de Oxford, Irena Watts, efectuó un descubrimiento de proporciones colosales. Cuando procedía a reparar la encuadernación de un salterio de los tiempos de Jacobo I, Watts se percató de que la cubierta del libro estaba formada por varias capas de papel atiesadas con engrudo. Consiguió separar las hojas sin destruir su contenido. Los papeles resultaron ser un capítulo desconocido hasta hoy de Los viajes de Gulliver, *la obra maestra de Jonathan Swift.* INVESTIGACIÓN Y CIENCIA *se ufana de publicar, con carácter de primicia, "Las nuevas aventuras de Gulliver en la isla volante de Laputa".*

En cosa de un mes había yo alcanzado una suficiencia tolerable en el lenguaje laputano, lo cual me permitió responder a casi todas las preguntas que el rey me hizo acerca del estado de las matemáticas en Europa. A modo de gesto graciable por su parte, el rey me invitó a visitar la Academia Volante Laputana, remedo de la institución mucho mayor que se asienta en tierra firme en la isla Balnibarbi. Me apresuré a aceptar, tanto por curiosidad como por cortesía, pues eran muchas las historias que sobre tal academia había oído.

Empezaron por presentarme a una especie de inventor que había consagrado 16 años a la construcción de un reloj de pared de doble péndulo —el mejor, según me aseguró su inventor, para saber la hora. Su intención primera había sido suspender el segundo péndulo de la luneta del primero y hacerlo de la manera más sencilla posible. Tras percatarse de que ciertas sutilezas de su teoría no podrían ser llevadas a la práctica, el inventor se había visto obligado a añadir muelles por aquí y contrapesos por allá, y a resultas, tras dieciséis años de modificaciones, la complejidad de la máquina había superado con mucho las previsiones iniciales. Al interesarme por la exactitud del reloj, se me respondió que señalaba la hora exacta dos veces al día.



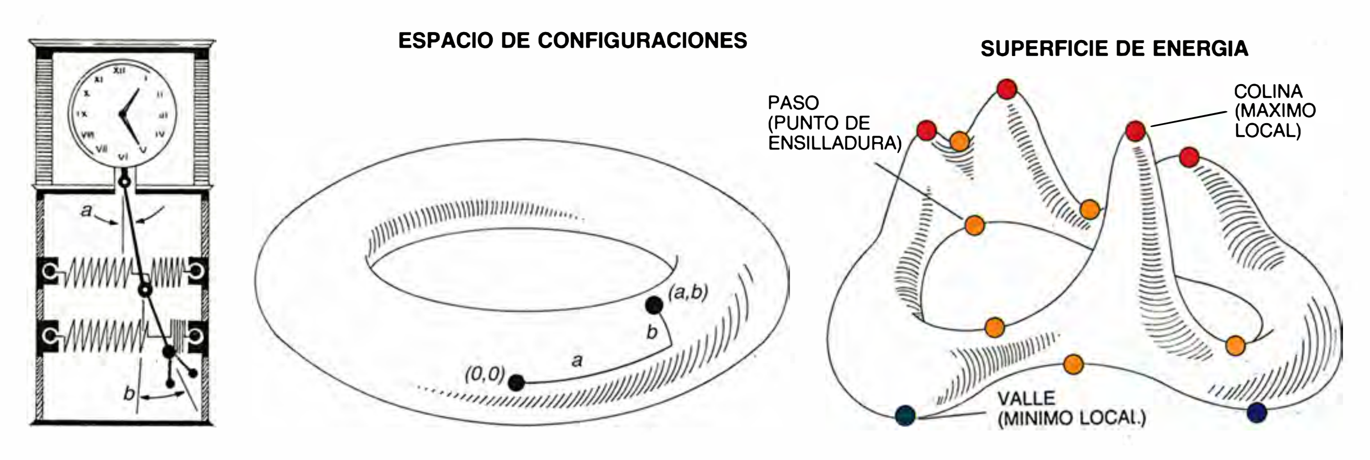
1. LA ISLA DE LAPUTA ha sido bendecida con 1267 colinas, 1506 valles y 1944 pasos, según el Real Servicio Topográfico. ¿Sabrá demostrar que el recuento es inexacto?

El relojero se mostró cordial en extremo cuando le felicité por la excelencia del fruto de sus trabajos y él me mostró los muchos cálculos matemáticos que determinaban el diseño de la máquina. Aunque soy incapaz de recordarlos todos, hay uno que perdura pertinaz en mi memoria. Para el correcto funcionamiento del reloj resultaba crucial que sus péndulos hubieran en ocasiones de estar en reposo, de forma que la acción combinada de los resortes produjera un equilibrio perfecto. La simplicidad del diseño original, me informó, hacía que tales cálculos resultasen a la vez sencillos y elegantes: el par de péndulos podía hallarse en equilibrio exactamente en cuatro posiciones.

Comprendí que, si ambos péndulos hubieran de colgar verticales y hacia abajo, permanecerían en reposo en esa posición indefinidamente. Sin embargo, empecé a poner en duda la existencia de otras tres de tales posiciones. Tras invertir mucho esfuerzo en ello, el relojero me hizo comprender que era posible una segunda, en la que ambos péndulos apuntaban verticalmente hacia arriba.

Admití que no se me había ocurrido una configuración semejante y una vez informado de que tal configuración podía, al menos en teoría, quedar en equilibrio, hice notar que, aun cuando un monje pudiera mantener en equilibrio una anguila sobre la punta de la nariz, tal conducta rara vez es observada ni en lonjas de pescado ni en monasterios. Pero el relojero insistió, describiendo tal disposición con el nombre de *thelmin frole*, que yo traduzco por "equilibrio inestable". En aquel momento creí oírle decir que la disposición era "de perversa inverosimilitud", sentir con el que me apresuré a estar de acuerdo. Habiendo entendido lo anterior, no tardé en deducir los dos estados de equilibrio restantes, en cada uno de los cuales uno de los péndulos queda en equilibrio apuntando hacia arriba y el otro pende verticalmente hacia abajo.

Lamentábase el inventor de que el único obstáculo que impedía llevar a término el proyecto consistiese en demostrar la existencia de cuatro o más configuraciones comparables en el artefacto real, resortes incluidos. Las posiciones exactas carecían de importancia; todo cuanto se requería era su existencia. Pero los problemas matemáticos estaban mostrándose inexpugnables y el relojero desesperaba de llegar jamás a hallar las soluciones. En ese trance me hallaba cuando fui rescatado por el mensajero del rey, quien me instó a pasar a las cocinas para un refrigerio de pudín de elip-



2. *EL MECANISMO del reloj* (izquierda) *se funda en dos péndulos. Los ángulos de los péndulos se corresponden con puntos de una superficie* (centro), *el espacio de configuraciones del reloj. Al llevar cada punto de la superficie hasta una altura igual a la energía de la configuración correspondiente queda definida la superficie de energía.*

soide y de añojo cónico, para proseguir después con la visita a la academia.

Tras el almuerzo me fue mostrada una cosa de la que el rey se hallaba muy ufano, un levantamiento geodésico de la totalidad de la isla. El topógrafo real era un personaje de faz rubicunda y cintura inmensurable, que siempre llevaba consigo una plomada a modo de emblema de su profesión. Su tarea, según me dijo, consistía en catalogar cada una de las colinas, valles y pasos de la isla.

Inquirí la definición precisa de tales términos, deseando aprehender la naturaleza exacta de la tarea. Por ejemplo, ¿se contaban como colinas los montículos de los hormigueros? El topógrafo respondió con orgullo que así era: una colina es cualquier prominencia cuya altura la hiciera sobresalir de su vecindad inmediata; era un flam, o sea, un "máximo local". Correspondientemente, un valle era un flim, un "mínimo local". El topógrafo real explicó que, técnicamente hablando, un paso era un lugar que tuviera un máximo local en una dirección pero un mínimo local en otra, algo así como una silla de montar. A decir verdad, tales lugares son conocidos por "puntos de ensilladura" en las grandes academias europeas, mas, obviamente, el término laputano para ellos no podía ser sino "flimflam".

La enumeración de estos accidentes, siguió explicando, había sido realizada con la más minuciosa y extrema precisión. Merced a este cálculo, aseguró sin modestia, existían en Laputa exactamente 1267 colinas y 1506 valles. Momento en el que yo interpuse que los pasos, consecuentemente, habrían de ser 2771.

El topógrafo afirmó que el Servicio Geodésico Real había cartografiado exactamente 1944. Yo repliqué que forzosamente se hubo de despreciar alguno, pues existe una relación entre el número de colinas, C, el de valles, V, y el de pasos, P. La relación es $C + V - P = 2$. La demostración es a la vez concluyente y general.

Informé al topógrafo real de que había cometido un error de 827 al contar el número de pasos, presumiendo siempre que hubiera efectuado un censo riguroso del número de colinas y de valles. Para apoyar el recuento oficial el topógrafo hizo desfilar por docenas ante mí a sus subordinados, quienes juraron la exactitud de sus observaciones. Pero poco a poco empezaron a surgir discrepancias y el topógrafo hubo de anunciar al poco que, debido a una pequeña omisión, el número de pasos tenía que incrementarse, convirtiéndose en 2772, persistiendo invariable el número de valles y colinas.

Aplaudí su diligencia, mas me aventuré a señalar que todavía subsistía una discrepancia de un paso. O bien había sobrestimado el número de pasos o había despreciado una colina o un valle. Me opuso que, por su naturaleza, es imposible desechar una colina, aunque admitió que, por igual motivo, se hubiera podido pasar por alto cualquier valle.

Entonces se exaltó mucho y me llevó ante otro de los miembros de la academia, un historiador. Confieso que jamás he visto un personaje más inverosímil para tal cargo, pues mal alcanzaba a recordar su propio nombre desde un segundo al siguiente. No obstante, a fuerza de gran empeño por su parte y de paciencia aún mayor por la mía, empezó a desarrollarse un remedo de narración.

Al parecer, no era yo el primer europeo que visitaba la isla de Laputa. Se rumoreaba que un tal capitán Kidd, de profesión pirata, había enterrado un tesoro en algún lugar de la isla "en el fondo de su valle más profundo". Tal tesoro no había aparecido a pesar de reiteradas búsquedas; tal vez mi afirmación de que se había olvidado un valle pudiera resolver el misterio. Era una lástima que mi método no facilitase la ubicación exacta del valle que faltaba.

Sin embargo, tras reflexionar nuevamente sobre el asunto me percaté de que "el valle más profundo" de Laputa no podría ser otro que el punto más bajo de la cara inferior de la isla, que era suave y redondeada como un plato. Unas discretas indagaciones me confirmaron que el Servicio Geodésico Real no había examinado la cara inferior de la isla y la discrepancia quedó resuelta a satisfacción mía. Me prometí inspeccionar por mi cuenta la parte más baja de Laputa.

Para distraer la atención pregunté si no podrían encontrarse en la isla arcos pétreos naturales. El topógrafo real me aseguró que existían, en efecto, algunos de ellos, aunque no sabía exactamente cuántos. Le informé de que el Servicio Geodésico forzosamente tenía que estar en un error. Mi demostración de que $C + V - P = 2$ se fundaba en la hipótesis de que no había agujeros en la isla. En toda superficie que posea orificios, como es el caso de los arcos pétreos, la relación entre los números C, V y P tiene que ser distinta de la que yo había deducido anteriormente. En nuestro caso, $C + V - P = 2 - 2g$, siendo g el número de agujeros.

Una inspección verdaderamente rigurosa, que determinase inequívocamente el número de colinas, valles y pasos permitiría deducir el número g de arcos pétreos. Por ejemplo, de haber 1000 colinas, 1000 valles y 2020 pasos, se verificaría la relación siguiente:

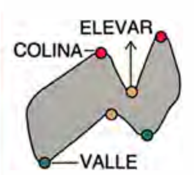
$$1000 + 1000 - 2020 = 2 - 2g$$

Teorema de Gulliver para superficies

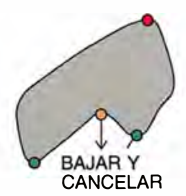
Cualquier superficie lisa y cerrada que contenga un número *C* de máximos locales, *V* de mínimos locales y *P* puntos de ensilladura habrá necesariamente de verificar la ecuación $C + V - P = 2$. La estrategia óptima para demostrar que así es consiste en deformar la superficie de forma continua, reduciendo estos números de modo tal que la expresión $C + V - P = 2$ se mantenga invariable.

La deformación consiste en una serie de movimientos. En cada uno de ellos, una colina o un valle se funde con un paso vecino, con lo que ambos desaparecen. El proceso continua hasta la eliminación de todos los pasos, tras lo cual solamente pueden subsistir una colina y un valle (pues de haber dos colinas o dos valles tendría que existir entre ellos un paso que los separase).

La fusión de una colina y un paso reduce en una unidad a cada uno de los números *C* y *P*, por lo que $C + V - P$ permanece invariable. La ilustración muestra el efecto de una secuencia de tales movimientos.

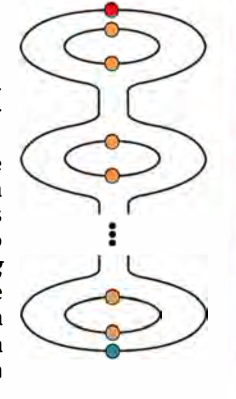








Una vez efectuadas todas las deformaciones, $C = 1$, $V = 1$ y $P = 0$, y por lo tanto $C + V - P = 1 + 1 - 0$, que es en este caso igual a 2. Dado que $C + V - P$ no varía a lo largo de toda la serie de deformaciones, el valor inicial de $C + V - P$ también ha de ser igual a 2.



La demostración de que $C + V - P = 2$ presupone que la superficie es topológicamente equivalente a una esfera. Sin embargo, una superficie general es equivalente a una esfera en la que se han perforado un número *g* de agujeros a su través. El valor de *g* determina el género de la superficie. Toda superficie lisa y cerrada de género *g* verifica necesariamente la ecuación $C + V - P = 2 - 2g$. La demostración es la misma que antes, pero la deformación concluye con la superficie que vemos a la derecha.

En este caso, $C = 1$, $V = 1$ y $P = 2g$ porque cada agujero genera dos pasos. Así pues, $C + V - P = 1 + 1 - 2g$, o lo que es igual, $2 - 2g$.

de donde resulta $2g = 22$, por lo que existirían exactamente 11 arcos pétreos. No me pareció que el topógrafo real se sintiera complacido por estas revelaciones; aun así, se comprometió a repetir el estudio topográfico con exactitud completa.

Mi pensamiento nunca se había alejado mucho del tesoro del capitán Kidd. Se me ocurrió que el pirata podría muy bien haber enterrado su tesoro en una profunda caverna llamada Flandona Gagnole, que me había sido mostrada ya en un momento anterior de mi visita. En el interior de la caverna se encuentra una piedra imán de tamaño prodigioso, montada sobre un eje. Mediante esta piedra imán se gobierna el ascenso o descenso de la isla y su desplazamiento a diferentes partes del mundo. Era posible que el tesoro se encontrase debajo de la piedra imán, por ser este lugar el punto accesible más próximo al más bajo de la isla. Decidí excavar en esa dirección. Por desgracia, antes de que mi túnel hubiera llegado a profundizar unos cuantos metros fui prendido y arrojado en prisión, acusado de sabotaje de la isla volante.

Permanecí allí tres días, cargado de cadenas; fui conducido luego ante el rey, quien me reconvino espaciosa y enérgicamente a causa de mi infracción de la ley laputana. Conseguí tras mucho porfiar que el rey se prestara a oír mis súplicas; le revelé mi deducción de que el gran tesoro pirata habría de encontrarse en la cara inferior de Laputa. El rey expresó su gratitud por mis ideas y acto seguido me condenó a cuatro días de trabajos mentales forzados.

Mi primera tarea, como era de suponer, consistió en hallar el modo de recobrar para el rey el tesoro enterrado. Resolví que lo más conveniente sería consultar con diversos miembros de la Academia Volante Laputana. Uno de sus filósofos afirmó que, si bien mi plan de perforar hacia abajo desde Flandona Gagnole era en principio perfectamente razonable, lo mejor sería realizarlo dándole la vuelta a la piedra imán, lo que provocaría que Laputa volase en posición invertida. Me apresuré a señalar que de hacerlo así todos los habitantes de la isla caerían de ella. Mi objeción fue rápidamente desechada por otro filósofo, que propuso que toda la población y sus posesiones podrían quedar firmemente asegurados aplicando una mano de cola fuerte sobre toda la isla.

Deseoso de evitar cuestiones tan pegajosas, concluí que la solución más sencilla consistiría en hacer descender la isla hasta unos pocos metros del suelo y erigir allí un andamio para examinar la cara inferior. Consideré la conveniencia de proponer yo mismo esta idea, mas recelé que podría ser yo el enviado a hacer la prueba y encontrarme con que la isla descendía por accidente sobre mi cabeza.

Quizás pudiese dar con alguien que transmitiera mi plan al rey. El relojero me pareció el candidato ideal, siendo personaje respetado por la academia. Accedió a presentar al rey mi propuesta, a condición de que le ayudase a demostrar que un sistema de dos péndulos tiene al menos cuatro estados de equilibrio, independientemente de cuantos resortes y contrapesos le sean acoplados.

Estuvimos de acuerdo en que las posibles configuraciones de la máquina están definidas por dos ángulos, los definidos por los dos péndulos. Estos dos ángulos se corresponden de manera natural con los puntos de una superficie específica, a saber, la superficie de un toro, a la que llamaré espacio de configuraciones de la máquina. Este toro tiene un agujero; por consiguiente su género es $g = 1$. Con independencia, pues, de cómo se encuentre situado el toro en el espacio, el número $C + V - P$ será siempre nulo.

La compresión de un muelle o la elevación de un peso conlleva el gasto de cantidades considerables de energía. Ello significa que cada configuración del mecanismo tiene asociada una magnitud, la energía total.

Imaginemos que cada punto de la superficie tórica es desplazado hasta una altura igual a la energía de la correspondiente configuración y llamemos superficie de energía a la superficie resultante. Los estados de equilibrio corresponden a posiciones de energía estacionaria, esto es, a colinas, valles y pasos de esta superficie.

Así pues, el número total de estados de equilibrio es $C + V + P$. Ahora, cada superficie finita ha de tener cuando menos un punto de altura máxima y, por tanto, al menos una colina, y, asimismo, al menos un punto de altura mínima y, por ende, un valle, de donde resulta que C es por lo menos 1 y V es al menos 1. Puesto que $C + V - P = 2(1) - 2 = 0$ se sigue que $P = C + V$ y, por tanto, P es al menos 2. Finalmente, el número de estados de equilibrio $C + V + P$ es al menos $1 + 1 + 2 = 4$, que es el resultado requerido por el relojero.

Mi método no especifica dónde se encuentran esos estados de equilibrio, pero sí tuvo éxito en proporcionar la cota inferior deseada para su número. Hago hincapié en que la conclusión es totalmente independiente de los resortes, cuerdas, pesas y demás lindezas: depende exclusivamente del número de agujeros del espacio de configuraciones. Me apresuré a explicar mi razonamiento al relojero y, tras considerable debate, mi demostración fue tenida por correcta, aunque extravagante.

El relojero cumplió su palabra y propuso al rey que se hiciera descender la isla volante hacia el suelo para efectuar una búsqueda del tesoro del pirata. El rey accedió y dio la orden. Antes de concluir el día, Laputa se cernía a escasos metros sobre el sólido suelo de la isla de Balnibarbi.

Mientras se preparaba una escala de cuerda para el aterrizaje, el topógrafo real solicitó audiencia al rey para informarle de que el segundo estudio geodésico real había quedado concluido. Los números, declaró con el ceño oscuramente fruncido en dirección a mí, eran 1893 colinas, 1942 valles y 3816 pasos. El rey pareció complacido al oír las nuevas. Sin embargo, según mis cálculos, el número g de arcos pétreos tenía forzosamente que venir dado por

$$2 - 2g = 1893 + 1942 - 3816 = 19$$

según lo cual, Laputa poseía $-8{,}5$ arcos.

Antes de que el rey pudiera percatarse de este hecho me despedí y me deslicé por la escala de garduño. A pesar de mi temor a verme aplastado, sentí la tentación de buscar en la parte inferior de la isla. Mas oyendo gran conmoción en la cara superior, partí y me alejé a toda prisa.

BIBLIOGRAFIA COMPLEMENTARIA

SURFACE TOPOLOGY. P. A. Firby y C. F. Gardiner. Chichester, Ellis Horwood, 1982.

Libros

Mecánica cuántica, Lipmann, grupos hiperbólicos, universidades e Iberia

L. Mas, M. Alemany, A. F. Costa, L. Alonso y D. Fletcher

LO DECIBLE Y LO INDECIBLE EN MECANICA CUANTICA, por J. S. Bell. Introducción y traducción a cargo de J. L. Sánchez Gómez. Alianza Universidad; Madrid, 1990.

Cuando un científico se dedica a estudiar temas que bordean la frontera o caen más allá de lo que la ciencia oficial (o de moda) llama ortodoxia es un firme candidato a la marginación más o menos expresada. No es éste el caso de J. S. Bell (1928-1990), investigador en física de partículas del CERN durante muchos años. Muy conocido, además, por las ahora llamadas "desigualdades de Bell" que desembocan en el "teorema de Bell", temas sobre los fundamentos lógicos de la mecánica cuántica.

Que la mecánica cuántica es una teoría aceptable desde el punto de vista fenomenológico, que da validez a las teorías, era y es una realidad observable y medible en muchísimos experimentos para los que, por ahora, no hay otro, y mejor debería ser, modelo físico.

Sin embargo, el aceptar y utilizar en el laboratorio una teoría física no es contradictorio con el estudio paralelo de sus fundamentos, lógicos y matemáticos. No sólo no es contradictorio sino necesario, y así se ha hecho siempre. Suele distinguirse entre el físico que se dedica a la parte experimental y el que se dedica a la parte teórica; muchas veces una misma persona abarca ambos aspectos, pero uno de los dos suele dominar claramente. En el caso de J. S. Bell, si consultamos su bibliografía, veremos que la distribución es bastante equilibrada.

Las controversias sobre los fundamentos de la mecánica cuántica vienen de lejos; siguen y seguirán. No debe engañarse el lector con la frase einsteiniana de que Dios no juega a los dados. Bell, en uno de los artículos del libro, cita otro pensamiento de Einstein más claro: "Estoy, de hecho, firmemente convencido de que el carácter esencialmente estadístico de la teoría cuántica contemporánea ha de achacarse únicamente al hecho de que ésta opera con una descripción incompleta del sistema físico". A él no le gustaba que un modelo físico basase sus resultados en una medición de probabilidades, pero reconocía que en aquel momento no había otro modelo, con verificación experimental, mejor. Por temperamento, o por voluntad de abrir brechas de investigación, buscaba otro modelo que mejorase la mecánica cuántica, modelo que debería basarse en la experiencia: de aquí la proposición del experimento conocido como de Einstein-Podolsky-Rosen, comentado en uno de los artículos del libro. Si se realizaba y daba el resultado por él esperado, aportaría una prueba experimental de que había "variables ocultas", más allá del cuadro de la mecánica cuántica, pero no contra el mismo.

Bell con sus "desigualdades" probó que Einstein tenía razón (lo acaba de recordar R. H. Romer), al suponer que habría una teoría "mejor" que englobaría la mecánica cuántica. Pero con su teorema demostró que Einstein no tenía razón (lo cuenta J. Gribbin), ya que del nuevo cuadro quedaría excluida la mecánica cuántica al proyectar una experiencia cuyo resultado estaría de acuerdo con su teoría general, pero no con el de la mecánica cuántica. Dicho experimento, y otros relacionados, se han realizado y parecen dar la razón a Bell. Sin embargo, la cuestión no está cerrada, como corresponde en física, y todo aquel que quiera participar y tenga la formación y medios intelectuales adecuados puede hacerlo.

Acaba de aparecer en *Physical Review Letters*, una de las revistas más prestigiosas y con el coeficiente de impacto que busca todo buen hacedor de currículum, un excelente y claro artículo en el que se demuestra la mala interpretación probabilística de uno de los experimentos citados. Si en vez de cociente de probabilidades nos limitamos a tomar probabilidades, el experimento estaría de acuerdo con las desigualdades de Bell y con la mecánica cuántica. Por cierto que el autor es E. Santos, de la Universidad de Cantabria, nuestro científico más conocido en este campo de los fundamentos de la mecánica cuántica, con una cantidad de trabajos sobre el tema considerable y tan interesantes como este último del *Physical Review Letters*.

El libro de Bell recoge todos sus artículos sobre el tema y algún otro afín, como el de "Cómo enseñar la relatividad especial". Los trabajos están ordenados cronológicamente, lo que considero, y así opina Bell en su introducción, no es el método adecuado. Sería mejor que hubiesen sido ordenados en grupos de subtemas, o bien grupos de dificultad creciente. Aun así, no sería fácil su lectura; una recopilación de artículos nunca lo es; cuando logras entrar en el ambiente del trabajo éste se acaba y en el siguiente, aunque sea del mismo autor, el ambiente y en parte el tema cambian. Una buena idea para la lectura es aceptar los consejos que da el mismo Bell en su introducción sobre distintas agrupaciones.

Algunos de los artículos parecen más digeribles, como el 15 sobre "mecánica cuántica para cosmólogos", que antes desconocía y en el que hay muchas ideas con las que coincido, y el 16 "Sobre los calcetines de Bertlmann y la naturaleza de la realidad". Se trata, quizás, de una impresión epidérmica, porque los títulos, llamativos, inducen fácilmente al engaño. R. Thom, medalla Fields de matemática, en una reciente entrevista, respondía así (a la pregunta "¿Debe ser el científico un poco filósofo?"): "Debería, pero no siempre es fácil. Hay una forma demagógica y popular de hacer filosofía que no es deseable". J. Bell no lo hizo así, sino que especuló sin concesiones a la galería. La dificultad del tema estriba en que para estudiar los fundamentos y la lógica de una teoría hay que saber, además de la teoría en sí, los fundamentos y la lógica de las ciencias en general.

Al interesado en el tema, y debería serlo en mayor o menor grado todo aquel que sepa o quiera saber mecánica cuántica, el libro le da una visión exacta de cómo ha ido evolucionando durante los años en los que J. S. Bell se ha interesado en su estudio. Si lo lee pausadamente y compara luego con otros puntos de vista, puede serle francamente útil. (L.M.)

THE ROOTS OF MODERN BIOCHEMISTRY. FRITZ LIPMANN'S SQUIGGLE AND ITS CONSEQUENCES. H. Kleinhauf, H. Von Döhren y L. Jaenicke (eds.). Walter de Gruyter; Berlín, 1988.

El texto que nos ocupa es un volumen dedicado a la memoria de Fritz Lipmann, uno de los más importantes científicos de nuestro siglo, que a lo largo de su dilatada carrera como investigador contribuyó de un modo esencial al desarrollo de la bioquímica. En casi mil páginas se recoge en merecido elogio de la figura de Lipmann lo más selecto de cuantos trabajan en el campo de la bioquímica y ciencias afines.

Muchos recordaremos a Lipmann por el símbolo ~ (virgulilla), que unido, normalmente a la P de fosfato: ~P, nos indica un enlace fosfato de alta energía. A ese detalle se refiere el subtítulo del libro, tomándolo como paradigma de las muchas contribuciones de Fritz Lipmann a la bioquímica moderna.

El libro tiene interés para los que quieran conocer las raíces de la bioquímica, una ciencia desarrollada casi enteramente en poco menos de setenta años. En este sentido, el libro ofrece numerosas impresiones personales de diversos autores que han convivido con Lipmann o que son sus coetáneos, así como panorámicas generales de los primeros tiempos del desarrollo de la bioquímica en Europa, principalmente en Alemania, y su posterior crecimiento expansivo al otro lado del Atlántico.

Lipmann, uno de sus más destacados protagonistas, nos es presentado como un hombre bueno, inteligente y tenaz trabajador, pero además tocado de intuición genial. En uno de los primeros capítulos, Christian de Duve resume magistralmente esta diferencia diciendo: "La mayoría de nosotros somos albañiles. Nos sentimos satisfechos con añadir algún ladrillo al edificio de la ciencia, y nos consideramos afortunados cuando contribuimos con una piedra angular, o la base de una columna, o la clave de un arco. Sólo unos pocos tienen la visión del arquitecto. Ven el edificio completo mucho antes de que esté acabado. Fritz Lipmann era uno de ésos." A esta visión la acompañó una personalidad extremadamente humana, cordial y generosa. No es, pues, de extrañar el gran número de contribuciones que forman este volumen dedicado a su memoria.



1. Fritz Lipmann (izquierda) *y Otto Meyerhof* (derecha). "*La fotografía es muy significativa. Nos tratamos muy poco y lo que yo aprendí de Meyerhof me vino por difusión, aunque ello habría de influirme para siempre. La fotografía se tomó en 1941, con ocasión de una reunión celebrada en Madison. Como puede apreciarse, se nos ve tensos. La razón es que yo había dictado una conferencia sobre el efecto Pasteur y me había mostrado discrepante de la interpretación que Otto daba.*"

El libro es muy desigual, poco organizado. Cada colaborador se mueve a su aire. Unos narran su relación con Lipmann y otros sacan a la luz retazos autobiográficos. Quienes han preferido contribuir a este homenaje con breves actualizaciones de campos específicos de la bioquímica o la biología molecular en los que su intervención ha sido destacada o fundamental. Hay otras contribuciones que desarrollan cortas revisiones sobre temas científicos mucho más especializados. Debemos reconocer el esfuerzo de los editores por establecer un mínimo de clasificación de los capítulos (1-Fritz Lipmann, 2-La mayoría de edad de la bioquímica, 3-La biología molecular afina sus herramientas, 4-La dinámica funcional y 5-Evolución), lo que facilita la consulta interesada. En conjunto el libro adolece de cierta disparidad de planteamientos (por ejemplo, hay artículos de muy diversa extensión, otros sin bibliografía y otros con abundantes referencias; la forma de presentar la bibliografía, además, difiere de un capítulo a otro). Se echa en falta un más cuidado trabajo de edición, que parece que se ha limitado en este caso a una breve clasificación y a la presentación conjunta de los textos con cortas introducciones a los grandes apartados del libro.

Los capítulos de recuerdo, biográficos y de matiz histórico, con ser sólo una parte relativamente pequeña del libro, revisten el mayor interés, ya que nos permiten reconocer la atmósfera pionera de los primeros laboratorios de bioquímica, en los que científicos de la talla de Lipmann ya veían con genial intuición el edificio de la moderna bioquímica.

Si bien muchas de las contribuciones del libro son de gran calidad, y pueden resultar útiles para el investigador científico, otras pecan bien de superficialidad o bien de superespecialización. En conjunto, el libro merece consultarse y su lectura (a trozos) es amena, en especial los dos primeros grupos de capítulos. Por su envergadura y objetivo declarado, sacarán el mayor partido los interesados en la historia de la bioquímica y los docentes de esta materia, ya que las cortas revisiones sobre temas puntuales les servirán para preparar algunos temas y seminarios. (M.A.)

GEOMETRIE ET THEORIE DES GROUPES. LES GROUPES HYPERBOLIQUES DE GROMOV, por M. Coornaert, T. Delzant y A. Papadopoulos. Springer Verlag; Berlín, 1990.

Los grupos hiperbólicos fueron de-

finidos por Mikhael Gromov en "Hyperbolic groups", *Essays in group theory*, editado por S. M. Gersten, M.S.R.I. (Springer-Verlag, 1987) páginas 75-263. Estos grupos forman una clase muy amplia de grupos de presentación finita donde se pueden aplicar métodos de origen geométrico para obtener propiedades similares a las que poseen los grupos fundamentales de variedades hiperbólicas. A título de ejemplo, como en el caso de los grupos fundamentales de variedades con curvatura negativa, se prueba que los grupos hiperbólicos no pueden contener subgrupos isomorfos a $Z \oplus Z$. Por otro lado los grupos hiperbólicos forman una categoría más amplia que dichos grupos fundamentales y, por ejemplo, el producto libre de dos grupos hiperbólicos es un grupo hiperbólico.

El artículo de Gromov era muy denso y duro. Ello explica la aparición posterior de varios libros donde se ofrecen versiones más detalladas y extensas de la teoría de grupos hiperbólicos. Entre dichos textos se encuentra *Geométrie et théorie des groupes. Les groupes hiperboliques de Gromov*. Otro libro donde se tratan los grupos hiperbólicos de Gromov es *Sur les groupes hyperboliques d'après Mikhael Gromov*, editado por E. Ghys y P. de la Harpe; también en las notas de B. Bowdich *Notes on Gromov's hyperbolicity criterion for path-metric spaces*, y *Notes on negatively curved groups*, editadas por H. Short.

Describamos sucintamente el contenido del libro firmado por Coornaert-Delzant-Papadopoulos. Comienzan definiendo espacio hiperbólico en el sentido de Gromov. Una definición geométrica de tales espacios hiperbólicos es la siguiente: un espacio métrico geodésico es hiperbólico si existe una constante $\delta \geqslant 0$ tal que, para cada triángulo geodésico, se verifica que cada lado está contenido en el δ-entorno de la unión de los otros dos lados.

A continuación se prueba que el espacio hiperbólico clásico es hiperbólico en el sentido de Gromov. Se define después el borde de un espacio hiperbólico en el sentido de Gromov y se estudian las cuasi-isometrías y cuasi-geodésicas de tales espacios. Más adelante se clasificarán las isometrías como en el caso de los espacios hiperbólicos clásicos en isometrías hiperbólicas, elípticas y parabólicas.

Se definen grupos hiperbólicos como los grupos finitamente presentados cuyos grafos de Cayley son espacios hiperbólicos en el sentido de Gromov. De forma semejante a lo que ocurre para los grupos fundamentales de variedades hiperbólicas, se construye, para un grupo hiperbólico dado, un complejo simplicial de dimensión finita, contractible y localmente finito sobre el que el grupo actúa con cociente compacto. Se caracterizan los espacios y grupos hiperbólicos mediante desigualdades isoperimétricas lineales (tal caracterización, que por supuesto ya está en el artículo de Gromov, no aparece probada en el texto de Ghys-de la Harpe, pero sí en las notas de Bowditch). Se estudia también la noción de cuasi-convexidad y sus aplicaciones a espacios hiperbólicos. Se desarrollan propiedades métricas del borde de los espacios hiperbólicos y se estudia la acción de las isometrías sobre el borde en relación con su métrica. Por último se encuadran los grupos hiperbólicos dentro de la clase más general de los grupos automáticos de J. Cannon y W. Thurston. (A.F.C.)

THE MEDIEVAL ENGLISH UNIVERSITIES. OXFORD AND CAMBRIDGE TO C. 1500, por Alan B. Cobban. The University of California Press; Berkeley, 1900. **NEW PERSPECTIVES ON RENAISSANCE THOUGHT.** Dirigido por John Henry y Sarah Hutton; Duckworth; Londres, 1989.

La historiografía de las universidades de Oxford y Cambridge se ha venido caracterizando por dos líneas paralelas de investigación: la historia de Oxford y la historia de Cambridge. Cobban ha roto esa orientación y ha creado un libro convergente; a imagen de la famosa regata que las enfrenta anualmente, la fundación, la organización, el desarrollo, la enseñanza, etcétera, de una se convierte en inmediato punto de referencia, de comparación para la otra. Podíamos generalizar aún más esa confluencia. Las Islas Británicas, por un lado, y el continente, por otro, han sido objeto de análisis, en cuanto a la institución universitaria se refiere, autónomos y cerrados. Carl B. Schmitt dedicó largos años de su vida, libros y artículos, a las universidades europeas, en particular las renacentistas y, de manera más determinada, las italianas de la Toscana. Sobre ellas se incide en el segundo volumen de la reseña, dedicado a la memoria de Schmitt. Tiene, pues, ante sí el lector dos herramientas poderosas para hacerse una idea de cómo eran las universidades europeas, desde el siglo XII hasta el XVI, de cuál era el conocimiento que en esos centros se transmitía, sus raíces y sus prolongaciones históricas; es decir, posee el marco general o el tapiz donde se teje la escena representada.

Aunque se hable de precedentes en las academias del mundo clásico grecorromano y bizantino, en la madrasa musulmana y en las escuelas catedralicias y palaciegas de la alta Edad Media, la verdad es que la aparición de las universidades en el XII supone una innovación radical. En un comienzo se trata de un movimiento espontáneo de gremialismo de profesores y alumnos, un "universo de gentes", una *universitas* que se quería defender y por eso se unía. Hasta el siglo XIV la institución docente veníase llamando *studium generale*. Y aun entonces la expresión aludía a la posibilidad de reunirse estudiantes de muy diverso origen o nación. Siendo plural el alumnado, una de las primeras notas definitorias de la universidad será el valor supranacional del grado obtenido en un centro. Como quiera que sólo el pontífice o el emperador tenían autoridad ecuménica, los centros, desde principios del XIII, se constituirán con una carta del primero o documento parejo del segundo: Federico II crea la de Nápoles en 1224, Gregorio IX la de Toulouse en 1229.

Existen documentos pertenecientes a las postrimerías del XII que reconocen la existencia del estudio general de Oxford *ex consuetudine*; esto es, venían dándose clases allí antes de la sanción papal, aunque con el apoyo del obispo, de quien dependía. Por costumbre funcionaron en esa centuria París, Bolonia, Montpellier y Reggio. Cambridge surge en torno a 1209, a raíz de la suspensión temporal de Oxford. El imaginarse ese período como una balsa de aceite político e ideológico se aleja mucho de la realidad. Quien ha estudiado, por ejemplo, ciencia árabe con Gerardo de Cremona en Toledo, aristotelismo en París y enseñado en Oxford tiene la mente bastante despejada para aborrecer provincianismos, en cualquier dominio: filosófico, político o jurídico. Si la intransigencia impide que en la Sorbona enseñen los mendicantes, Oxford les abrirá las puertas de par en par; si allí se prohíbe la física de Aristóteles, se comentarán aquí sus "libros naturales".

Cambridge aporta a la historia de las universidades el privilegio o *ius non trahi extra*, que le concedió Gregorio IX en 1233. Significaba que los alumnos o profesores no podían ser juzgados de delito fuera de las lindes de la diócesis de Ely, a la que pertenecía la universidad; en particular, sólo estaban sujetos a la autoridad del

obispo y del canciller. Ese privilegio de foro junto con el *ius ubique docendi* (reconocimiento universal de los grados) constituirán los quicios sobre los que se asiente la autonomía universitaria y, en definitiva, la persistencia de la institución.

París se convirtió en el modelo de la organización y gobierno de las universidad del norte de Europa, y, por consiguiente, de Oxford y Cambridge. Modelo que podríamos simplificar en una estructura jerárquica en oposición a la más democrática, hasta cierto punto, del mediodía continental. Así, en línea de principio, el canciller reúne la plenitud de poderes, primero por delegación del obispo y siglos más tarde, conseguida mayor autonomía, por razón del cargo. En ausencia del canciller, una situación que comenzó a hacerse frecuente desde mediados del siglo XV, le sustituía el vicecanciller, auxiliado por los *hebdomadarii*, una suerte de jefes de negociado que rotaban por semanas y estaban encargados de despachar los asuntos menores. Al frente de todos los asuntos administrativos estaba el "proctor", que, en rango, venía detrás mismo del canciller. La figura por excelencia de la universidad europea, de Bolonia a Oxford, era la del bedel. No era la suya una responsabilidad menor: recogía las tasas y cobraba las multas, era jefe de ceremonias, cuidaba del funcionamiento normal, vigilaba la limpieza; a veces eran propietarios de los contornos y servían de puente entre la universidad y la ciudad. Tal importancia revistieron que muchos monarcas no dudaron en intervenir en la elección de los bedeles (Enrique VI, Margarita de Anjou, Eduardo VI y Enrique VII). La institución contaba también con bibliotecario, registrador, proveedor, etc.

Aunque gozaron de favores reales y de la protección de la Iglesia no nadaban las universidades inglesas en la abundancia. Malvivían de las rentas de donaciones y de las tasas que cobraban por la colación de grados. Lo que no fue óbice para que, lo mismo Oxford que Cambridge, cuidaran de su biblioteca, cuyos fondos de manuscritos griegos y latinos empezó muy pronto a crecer.

No podemos separar la universidad de los colegios mayores. La verdad es que irían éstos adquiriendo tal importancia que, en el XVI, podemos ver ya el alma mater como una federación de esos institutos. Nacidos como obra pía para subvenir las necesidades de los estudiantes, poco a poco irían elevando el tono del cuadrivio (aritmética, geometría, música y astronomía)



2. Eduardo III concede privilegios reales a la Universidad de Oxford en 1375 (Biblioteca Bodleyana).

para convertirse en una suerte de facultades menores donde se enseñaba lógica, filosofía natural y lenguas clásicas. Del siglo XIII y XIV datan nombres que habrían de expandir la fama de Oxford y Cambridge: Merton, King, Trinity, Corpus Christi. No tardarán en fundarse Queens, All Souls, Jesus, Magdalen. Ese movimiento, aunque de origen continental —piénsese en el parisiense Colegio de Navarra, donde enseñó Nicolás de Oresme—, tan sólo se consolidaría a lo largo de los siglos en las universidades inglesas.

Se afirma con ligereza que la universidad no investigaba y que se limitaba a repetir, por vía de comentario, lo dicho por el autor seguido, sea Aristóteles, sea Galeno. Basta la remisión a la observación de los naturalistas, al juego de espejos de los

perspectivistas o a la anatomía comparada de ciertos médicos, para negarlo. Para su ventura, lo que no había era distinción entre ciencias y letras. Sí presentaba mayor viveza el desarrollo del curso, con sus clases lectivas ordinarias y clases extraordinarias, sus disputas públicas y la *determinatio* a modo de resumen sentenciador. Posee Cobban especial habilidad para conjugar la explicación llana del desarrollo de un curso, de un programa académico, del nombramiento de maestros, de la relación subordinada entre colegios, de dar vida a esa página sepia del pasado en definitiva, con un aparato documental mínimo, aunque suficiente. De todas formas, como dijimos más arriba, este es el marco.

Schmitt fue, por encima de todo, un historiador de las ideas. En éstas se centra *New perspectives*..., a él dedicado y que se abre con un ensayo, de Richard H. Popkin sobre una de las notas características del humanismo renacentista: el escepticismo, considerado aquí desde la óptica del enfrentamiento dialéctico entre judíos y cristianos. El escepticismo filosófico encuentra su expresión máxima en el antiaristotelismo de Pico della Mirandola y en nuestro Francisco Sánchez, pero no es esa toda la verdad, viene a escribir Eckhard Kessler, para quien la doctrina del estagirita sigue, adaptada, vigente en neoplatónicos y humanistas, amén de constituir una corriente autónoma.

Traza las coordenadas del aristotelismo Richard Sorabji, en particular en las modificaciones que sufre, a propósito del tema de la creación y persistencia del mundo, con sus temas colaterales de dinámica (principio de inercia, vacío y falta de rozamiento que posibilita la rotación incesante). Hilvana la historia arrancando de Proclo y Amonio, para aclararnos la posición de Filopón y, más tarde, de los árabes y la escolástica. John Murdoch insiste en el carácter ecléctico del neoaristotelismo renacentista que se plasma en el reconocimiento de otras autoridades, la incorporación de nuevos temas y nuevo enfoque del movimiento. Es el neoaristotelismo una postura abierta, capaz de admitir cuanto de verdad se diga, de acuerdo con el viejo aforismo de la Escuela según el cual Aristóteles no era racional por su autoridad, sino que la autoridad le venía de su racionalidad. Lo que se puso de manifiesto en la incorporación del magnetismo de William Gilbert, teoría contraria a la física aristotélica embalsamada, como expone aquí Stephen Pumfrey.

Otro grupo de artículos están dedicados a reseñar la figura de algún personaje central del Renacimiento. Así el de Charles Webster sobre las reservas de Gessner ante la oleada de paracelsismo que, so capa de nueva farmacología, amenazaba la fe protestante; también el de Thomas B. Settle sobre Egnazio Danti, un dominico de la corte de Cósimo que pintó, esculpió y forjó el conocimiento matemático, geográfico y astronómico de su tiempo; o el de M.J.B. Allen, experto en neoplatonismo, que es lo mismo que decir en Marsilio Ficino, a cuya obra introductora de la corriente hermética dedica su ensayo; o, por fin, el de Vivian Nutton, sobre Antonio Musa Brasavola y sus curiosos prólogos áulicos. No se agota, empero, en eso la riqueza de este volumen que trae, en cada página, la huella o la sombra de un maestro de su disciplina; Schmitt, por supuesto. (L.A.)

LEXICO DE INSCRIPCIONES IBERICAS (1976-1989), por Javier Velaza. Universidad de Barcelona; Barcelona, 1991.

El libro que ahora comentamos es la tesis doctoral defendida por su autor, Javier Velaza, en la Universidad de Barcelona. Puede considerarse como la continuación de otra tesis con el mismo título leída por Jaime Siles en 1976, aunque publicada en 1985, razón por la cual quedó fuera de su estudio la rica floración de textos ibéricos (Los Villares, Pico de los Ajos, Orleyl, Enguera, Palamós, Pech Maho, etc.) que han venido a enriquecer el acervo léxical ibérico en grado tal que nos proporcionan nuevas perspectivas, permitiéndonos alcanzar firmes logros dejando de ser muchas cuestiones meras hipótesis de trabajo para convertirse en realidades.

Y son estos textos, dados a conocer en publicaciones aparecidas entre 1976 y 1989, los recogidos en el libro de Velaza; de ahí el por qué de su título delimitando fechas, que sirven para diferenciar ambos léxicos.

Para los textos publicados con posterioridad a 1989 sugiere el profesor Untermann, autor del prólogo del libro, la necesidad de ir preparando "un tercer léxico (Léxico Ibérico 1899-1999)" para completar el Corpus de textos ibéricos hasta fines del presente siglo.

La aparición del "Léxico Velaza" viene casi a coincidir con la del vol. III de los "Monumenta Linguarum Hispanicarum" (Wiesbaden 1990) del citado profesor Untermann, y aunque prácticamente los textos del Léxico se hallan recogidos en los Monumenta, el libro de Velaza ofrece la ventaja de su cómodo manejo, estar redactado en español y su coste asequible al estudioso. Por el contrario, creemos que en el Léxico debieron reproducirse, junto a su lectura en alfabeto actual, los textos originales, lo que no entrañaba dificultad alguna puesto que ya estaban publicados. Al lector le hubieran evitado tener que buscar la parte gráfica en otras publicaciones.

En la Introducción (p. 3/12) el autor explica con todo detalle el método seguido en la confección de las fichas que reseña en páginas siguientes. Cada voz ibérica va precedida de un número de orden; sigue su lectura indicándose, si las hay, las discrepancias con otros autores, los paralelos y comentarios, el soporte en que descansan los letreros (plomo, cerámica, piedra, hueso, etc.), lugar donde se conserva el original y finaliza con la oportuna bibliografía.

Tal vez, tras el correspondiente número de orden y la voz ibérica, hubiera ido mejor la denominación y numeración que dieran los autores que la publicaron, sin necesidad de utilizar como identificación base la clave Untermann, que en la realidad es menos práctica que la empleada por los iberistas españoles.

Destina la p. 23 a las abreviaturas de los museos en donde se hallan parte de los textos utilizados en el Léxico. Una pequeña observación a este respecto, las siglas MPDV y SIP se refieren a un mismo organismo, por lo que, en futuras ediciones, deberán unificarse.

Al repertorio de voces, parte básica del estudio, están dedicadas las p. 24 a 179, numerándose del 1 al 890 (en realidad al 881), debido a un salto en la numeración (878/888). Constituye una rica aportación de datos, por lo que será de obligada consulta en todo futuro trabajo sobre estos temas.

Estas 881 entradas proceden de más de cincuenta yacimientos, de los que el autor transcribe (p. 180/189) dieciocho textos completos: Caspe; Orleyl V, VI, VII y VIII; Palamós; Pontós; Tarragona I y II; Los Villares VI y VII; Yátova (su denominación correcta Pico de los Ajos) I, II y III; Pech Maho I, II, III y IV.

Termina la obra con la relación de voces citadas en el texto y las variantes de lectura. Los números que aparecen entre paréntesis deberán adecuarse correctamente con los de las fichas, pues en ocasiones no coinciden.

El libro es de indudable utilidad para los estudiosos del tema y, si añadimos lo cómodo de manejar, no dudamos que será consultado con asiduidad, lo que obligará a una segunda edición, cosa que celebraremos sea a la mayor brevedad. (D.F.)

Apuntes

Los basaltos del rosario de islas del hemisferio austral portan una signatura distintiva de alta concentración de iones plomo y estroncio. La signatura Dupal —así se llama en honor de dos geólogos franceses, Dupré y Allègre— se origina en el manto profundo; según acaba de medirse, a unos 670 km por debajo de la frontera entre el manto superior y el inferior.

Los exones se hallan interrumpidos por los intrones: segmentos de ADN carentes de función codificadora y desechados en la transcripción de los genes de ARN. Se descubrieron en 1977 y desde entonces su origen no ha dejado de inquietar a los investigadores. ¿Llegaron en un momento precoz de la evolución para ensamblar los fragmentos sueltos del primer genoma? ¿Son de adquisición tardía? Las pruebas más sólidas favorecían la segunda opción: sólo se habían descubierto en animales y vegetales, pero no en las eubacterias; éstas, con archibacterias y eucariotas, comparten un antepasado común. Acaba de evidenciarse su presencia en cianobacterias, lo que constituye un balón de oxígeno para la hipótesis que defiende la remota antigüedad de los intrones, si bien no resuelve la cuestión de manera definitiva. ¿Y si las eubacterias los hubieran adquirido al correr el tiempo?

Si abrimos un manual de ecología, y buscamos la definición de mimetismo batesiano, leeremos algo parecido a lo siguiente: ''Llamado así en honor de H. W. Bates, consiste en que especies indiferentes e inofensivas, que adoptan el mismo uniforme de las aposemáticas (que disuaden por su color, toxicidad, etcétera), pueden gozar de cierta protección''. El ejemplo clásico, aunque no único, era el de dos géneros (*Limenitis* y *Danaus*). Habrá que ir borrándolo, o convertirlo, según parece, en muestra de mimetismo de Müller: en virtud del cual los dos mimetas, igualmente aposemáticos, se protegen para mutuo provecho.

Sabido es que la disponibilidad de agua limita la instalación permanente de una población. Tikal, en Guatemala, fue un asentamiento maya que alcanzó su máximo explendor a mediados del siglo octavo y fue abandonado en torno al año 1000. Alejado de toda fuente natural de agua capaz de abastecer la ciudad, de 60.000 a 80.000 habitantes, sufría cuatro meses de sequía anual. Para asegurar su suministro creó una compleja red de depósitos (bajos, pozas y aguadas) cuya gestión se repartían las clases pudientes. Mientras se buscan razones para explicar el hundimiento maya, comienza a apuntarse la imposibilidad de asegurar ese recurso imprescindible para la vida de los ciudadanos y el desarrollo de muchas actividades.

Para acercarse a la filosofía presocrática los alumnos saben que deben manejar el ''Kirk-Raven'', un compendio clásico cuya primera edición apareció en 1957. Lo que la inmensa mayoría desconoce es que John Raven (1914-1980) era un experto naturalista, capaz de corregir las especies erróneas de la *Historia plantarum* de Teofrasto publicada por la Loeb. Ello le permitió moverse con especial soltura entre los *physiologoi* que precedieron a Aristóteles.

Con la apertura del canal de Suez en 1869 se juntaron dos áreas zoogeográficas muy distintas: el mar Mediterráneo subtropical, que se conecta con el Atlántico por Gibraltar, y el mar Rojo tropical, la extensión septentrional de oceano Indico. El paso unidireccional y el avance de las especies desde el Rojo al Mediterráneo se conoce por migración Lesseps, en honor del ingeniero del canal. Hasta ahora se supone que han cruzado ese estrecho unas 200 especies.

La focalización por gravitación podría ayudar a quienes buscan enanas marrones donde pudiera yacer buena parte de la masa de nuestra galaxia. Esos objetos subestelares, muy pequeños y postulados por hipótesis, no pueden verse por su propia luz. Ahora bien, la masa central de la galaxia está densamente poblada de estrellas de movimiento lento; al pasar éstas por detrás de estrellas que, en el disco galáctico, nos son más cercanas, la gravedad podría enfocar su luz y darnos una imagen brillante. Nada impide que las enanas pardas se presten a ese juego con las estrellas más distantes, si bien el brillo sería de menor duración.

Se ha repetido por todos los medios el elenco de males que ha traído la guerra del golfo Pérsico para la biosfera: contaminación ambiental por combustión de yacimientos que tardarán años en apagarse, contaminación química por bombardeo de fábricas iraquíes y pérdida de estuarios y ecosistemas marinos por el vertido de petróleo al mar. De lo que no se ha hablado es de la ''contaminación geológica'': la alteración del desierto que limita a los países implicados (Arabia Saudita, Kuwait e Iraq) provocará un incremento espectacular de las tormentas de polvo y arena, que originará nuevos arenales y nuevas dunas que podrían cubrir aeropuertos, campos agrícolas experimentales y ciudades enteras.

Seguiremos explorando los campos del conocimiento



TERAPIA DE CHOQUE, por Justin A. Zivin y Dennis W. Choi
Aunque todavía no existe un tratamiento eficaz para paliar las lesiones cerebrales, en el horizonte aparecen muchas opciones prometedoras. Algunas de ellas ya se están evaluando en importantes ensayos clínicos con seres humanos.

INTERFEROMETRIA OPTICA DE SUPERFICIES, por Glen M. Robinson, David M. Perry y Richard W. Peterson
Basándose en la naturaleza ondulatoria de la luz y en la potencia de los ordenadores modernos, los autores de este artículo han diseñado unos dispositivos sensibles capaces de medir la textura de las superficies.

CONTROL BIOLOGICO DE MALAS HIERBAS, por Gary A. Strobel
Insectos y microorganismos sirven ya como herbicidas comerciales, y otros enfoques biológicos resultan prometedores. El objetivo: alternativas ambientalmente compatibles a los herbicidas químicos.

METALURGIA DE ALEACIONES DE COBRE EN EL PERU ANTIGUO, por Izumi Shimada y John F. Merkel
Hace más de mil años, la civilización de América del Sur se basaba en aleaciones de cobre. Recientes excavaciones han sacado a la luz muchos aspectos de la metalurgia de dichas aleaciones, desde la minería hasta su fusión y posterior conformación.

COMPLEJO DE CABO ORTEGAL, por J. I. Gil Ibarguchi y Jacques Girardeau
Restos de corteza oceánica, secciones de zonas orogénicas profundas y fragmentos de manto superior terrestre aparecen en varias unidades estructurales de un complejo geológico en la región del Cabo Ortegal.

LA VIDA PRIMITIVA DE LAS ESTRELLAS, por Steven W. Stahler
Las violentas colisiones de gas dominaron la infancia de las estrellas. La fusión nuclear permitió luego a las estrellas maduras arder durante miles de millones de años.

DISPERSION AUSTRONESIA Y ORIGEN DEL LENGUAJE, por Peter Bellwood
Los lenguajes austronesios del Pacífico se propagaron por tierra y mar a lo largo de 1500 años. La iniciativa partió de la necesidad de nueva tierra por colonizar.

¿LISTOS PARA VOLAR?, por Gary Stix
La modernización del transporte guarda relación directa con la incorporación de los sistemas automáticos. El problema reside en la conjunción inteligente entre el hombre y la máquina.

